Public's Inflation Expectations and Monetary Policy

Leonardo Melosi[†]

London Business School

28 February, 2011

Abstract

The paper develops a DSGE model where monetary policy propagates by affecting and coordinating price setters' expectations. Price setters face costs of price adjustment and do not observe the history of technology, monetary, and preference shocks. They form expectations about the evolution of their nominal marginal costs observing their idiosyncratic productivity, which is correlated with the aggregate productivity, last period's output and inflation, and the interest rate set by the central bank according to a Taylor rule. The model is estimated through likelihood methods on a U.S. data set including the Survey of Professional Forecasters as a measure of price setters' expectations. The transmission channel based on price-setters' expectations is found to account for the following three empirical facts that have been documented by the VAR literature: (*i*) the delayed and persistent effects of monetary shocks on inflation with associated sluggish response of real output; (*ii*) the price puzzle; (*iii*) the disappearance of the price puzzle after the 1970s. Structural policies of disinflation, such as the introduction of an "inflation targeting", need to be backed up by a perfect commitment device to succeed.

Keywords: Imperfect common knowlegde, heterogenous beliefs, price puzzle, persistent real effects of nominal shocks, Bayesian econometrics.

JEL classification: E5, C11, D8.

[†]Correspondence to: Leonardo Melosi, London Business School, Regent's Park, London, NW1 4S, United Kingdom. Email: lmelosi@london.edu. I wish to thank Harald Cole, Matthieu Darracq Pariès, Jesus Fernandez-Villaverde, Dirk Krueger, Kristoffer Nimark, Massimo Rostagno, Frank Schorfheide, and Paolo Surico for their very helpful comments and discussion. I also thank the participants of the joint ECB/Bundesbank/University of Frankfurt seminar for their very helpful comments. I thank the European Central Bank for its hospitality during part of this project.

1 Introduction

That monetary policy influences output and inflation by affecting public' expectations has come to a growing consensus among scholars and policy makers[1] in the last twenty years. Woodford (2005) writes:

> *Because the key decisionmakers in an economy are forward-looking, central banks affect the economy as much through their influence on expectations as through any direct, mechanical effects of central bank trading in the market for overnight cash.*

As pointed out by Morris and Shin (2008), market participants take decisions based on variables that are out of direct control of the central bank but they look at central bank's actions for clues about where those variables are headed. In this view, public's expectations are dispersed and can be coordinated by publicly observed action, such as the interest rate set by the central bank. The literature has produced a few stylized models where the central bank can coordinate dispersed agents' expectations in self-fulfilling environments (e.g., Morris and Shin, 2003 and Hellwig, 2002). On one hand, these models are too stylized to be well-suited for a formal investigation about the empirical significance of this new propagation channel for monetary impulses. On the other hand, the workhorse models (e.g., Christiano, Eichenbaum, and Evans 2005 and Smets and Wouters, 2007) for evaluating how monetary shocks propagate do not feature any explicit role for agent's expectations. A standard assumption in these models is, in fact, that agents have complete information, that is, they observe the complete history of the shocks that have hit the economy. As a consequence, agents share the same expectations about the model variables (e.g., GDP, inflation, interest rate, etc.). Furthermore, agents' expectations correctly forecast the evolution of model variables in the aftermath of an unanticipated monetary policy shock. Therefore, in these models there is no scope for the central bank to manage and coordinate such expectations by setting the interest rate or making announcements.

The paper bridges the gap in the literature by constructing and estimating a Dynamic Stochastic General Equilibrium (DSGE) model in which monetary policy can influence the macroeconomic aggregates (e.g., GDP and inflation) via affecting and coordinating dispersed agents' expectations. The model features technology, monetary, and preference shocks. Price setters observe their idiosyncratic technology shocks (i.e., a *private* signal), which conveys information about the aggregate productivity. Furthermore, the price-setters observe last period's real output and inflation as well as the interest rate, which is set by the central bank

[1] See, for instance, Bernanke (2004) and King (2005), and Trichet (2006).

according to a Taylor rule (i.e., a *public* signal). This signal provides information about the current inflation and output to the price setters. Price setters face cost of price adjustment in the form of Calvo sticky prices (Calvo, 1983). Hence, those price setters who are allowed to optimize their prices need to forecast the evolution of their nominal marginal costs when taking their price-setting decisions. In such an environment, firms, hence, face higher-order uncertainty as the optimal price will depend not only on their beliefs about the future price levels, but also on their beliefs about other price setters' beliefs about the future price levels, etc. (Townsend 1983a and 1983b).

It is important to notice that the monetary policy rule (i.e., the Taylor rule) plays a twofold role in the model. First, it is one of the transition equations for the model endogenous state variables (i.e., inflation, output, interest rate). Second, this reaction function is the stochastic process that drives the public signal of the policy rate. This second role adds a new transmission channel for monetary policy to the otherwise standard New Keynesian model. This new channel is based on the responses of price setters' expectations to monetary impulses and emerges for the following three reasons. First, price setters have to form expectations about the evolution of their nominal marginal costs because they face cost of price adjustment. Second, all price setters perfectly observe the current policy rate and this is common knowledge among them. Third, the central bank follows a monetary policy rule with feedbacks, that is, it reacts to endogenous variables, namely current inflation and real output. When most of the variability of the interest rate stems from these endogenous variables rather than from exogenous variables (i.e., monetary shocks), then a rising policy rate will be interpreted by price setters as evidence of growing inflation and rising real marginal costs. Consequently, a monetary shock will lead price setters to mistakenly believe, in the short run, that the central bank is reacting to non-monetary shocks (i.e., technology or preference shocks). As the paper shall show, adding this transmission channel based on affecting price setters' beliefs to an otherwise standard New Keynesian model ends up enriching the dynamic effects of monetary impulses.

The model is estimated through likelihood methods on a U.S. data set that includes the *Survey of Professional Forecasters* as a measure of price setters' inflation expectations. The data range includes the 1970s, which were characterized by one of the most notorious episode of substantial rise in inflation and inflation expectations in recent US economic history. The paper finds that the model fits the observed inflation expectations very well. Furthermore, the transmission channel for monetary policy based on affecting price setters' expectations is found to be empirically relevant, accounting for three important facts that the VAR literature shows to characterize the dynamic effects of monetary shocks: (i) the inflation persistence puzzle (i.e., a slow and delayed fall in inflation in response to a contractionary monetary policy shock)[2], (*ii*) the price puzzle (i.e., a temporary rise in the price level after a contractionary monetary policy shock)[3], (*iii*) the disappearance of the price puzzle after the 1970s.[4]

The estimated model predicts that a rise in the policy rate due to an unanticipated monetary shock is, at first, interpreted by price setters as a response of the central bank to a positive preference shock. Hence, a monetary tightening raises price setters' expectations about future real marginal costs, which in turn will drag current inflation upwards in the short run. The channel based on price setters' expectations thus makes the impact of a monetary shock upon inflation more delayed and persistent. This feature allows the model to capture the fact (*i*). The short-run effects of a monetary tightening upon price setters' expectations are so strong that the model can explain the price puzzle, that is, the fact (*ii*). When the paper estimates the model on a subsample that does not include data on 1970s, it finds that the price puzzle disappears after the 1970s, that is, the fact (*iii*).

The paper also uses the model to study the effectiveness of structural policies of disinflation (e.g., reduction of the inflation target for monetary policy or the introduction of an *inflation targeting*). I find that the presence of a device that helps coordinating price setters' inflation expectations (e.g., a viable commitment device) is key for such policies to succeed with very limited costs in terms of output loss. The paper finds that structural policies of disinflation that are not backed up by a viable commitment device (i.e., price setters have to learn the new inflation target by observing the dynamics of the policy rate) are doomed to fail. Five years after the implementation of a policy that reduces the inflation target by 50 basis points (bps), the inflation is predicted to go down by only 6 bps. Furthermore, the cost of disinflation associated with this policy would be 10.8 times larger than when price setters can immediately learn the new inflation target (e.g., perfect commitment). The reason why structural policies of disinflation fail under no commitment is that they are considered as rare events by price-setters. Therefore, price setters learn about such regime changes at a very slow pace. Consequently, sluggish adjustments of price setters' inflation expectations lead the new policy regime to fail reducing inflation in the long run. Moreover, even though a central bank cannot perfectly commit, being very aggressive against inflation is found to make the new regime to be effective in reducing inflation in the long run. The reason behind this result is that strong central bank's responses to inflation turn out to speed up price setters' learning about the new inflation target. Nonetheless, this positive outcome comes at the cost of very large output losses.

[2]See Christiano, Eichenbaum, and Evans (1999) and Stock and Watson (2001).

[3]See Sims (1992) and Eichenbaum (1992).

[4]See Barth and Ramey (2001), Hansen (2004), Castelnuovo and Surico (2010). and Ravn, Schmitt-Grohe, Uribe, and Uusküla (2010).

2 A Brief Overview of the Literature

From a theoretical perspective, the idea that publicly observed policy can coordinate agents' expectations has been recently explored by the literature of global games (Morris and Shin, 2003a). Morris and Shin (2003b) and Amato and Shin (2003, 2006) derive normative implications for incomplete-information settings and focus on the welfare effects of disclosing public information. Hellwig (2002) derives impulse responses to a large range of shocks for an economy with monopolistic competition and incomplete information. These partial equilibrium models, however, are too stylized to be used for empirically assessing central banks' role for coordinating expectations.

This paper is closely related to Nimark (2008), who introduces a model where firms hold private information about the dynamics of their future marginal costs, and face both strategic complementarities in price setting and nominal rigidities. The nice feature of this model is that the supply side of this economy can be analytically worked out and turns out to be characterized by an equation that resembles the standard New-Keynesian Phillips curve. Nonetheless, unlike the present paper the role of monetary policy in coordinating agents' expectations is absent in that central bank's actions only convey redundant information to agents.

The model that is presented in this chapter is also related to Lorenzoni (2010), who studies optimal monetary policy in a price-setting model where aggregate fluctuations are driven by the private sector's uncertainty about the economy's fundamentals. The main difference from my paper is the imperfect observability of the policy rate. The key mechanism of my paper is based on the fact that the policy rate is a public signal that conveys non-redundant information to price setters. In Lorenzoni's paper, this mechanism is absent in this paper and monetary policy rules matter only because they affect agents' incentives to respond to private and public signals.

The paper is also related to the literature that uses incomplete-information models for studying the persistence in economic fluctuations (Townsend, 1983a, 1983b; Hellwig, 2002; Adam, 2009; Angeletos and La'O, 2009; Rondina, 2008; and Lorenzoni, 2009) and the propagation of monetary disturbances to real variables and prices (Phelps, 1970; Lucas, 1972; Woodford, 2002; Adam, 2007; Gorodnichenko, 2008; Nimark, 2008; and Lorenzoni, 2010).[5]

The paper relates to the quickly growing literature that studies rational inattention models. The space in this section is regrettably too small to do justice to all of them. The closest studies to this paper are Woodford (2002), Maćkowiak and Wiederholt (2009), Maćkowiak

[5]See Mankiw and Reis (2002a, 2002b, 2006, 2007), and Reis (2006a, 2006b, 2009) for models with information frictions that do not feature imperfect common knowledge but can generate sizeable persistence.

and Wiederholt (2010), and Paciello (2010) who calibrate models in which firms have limited attention to match few stylized facts about lasting real effects of monetary disturbances. Maćkowiak, Moench, and Wiederholt (2009) and Woodford (2009) develop rational inattention models that account for substantial real effects of nominal disturbances and evidence on the frequency of price changes at micro level. Melosi (2010) performs an econometric analysis of these type of models which delivers the following results. First, these models - albeit very stylized - can capture the persistent and hump-shaped response of output and inflation to nominal disturbances observed in the data. Second, limited attention is better suited than price stickiness à la Calvo (1983) to account for the highly persistent effects of nominal shocks.

3 The Model

The economy is populated by perfectly competitive final-good producers (or, more briefly, the producers), a continuum $(0,1)$ of intermediate-good firms (or, more briefly, the firms), a continuum $(0,1)$ of households, a central bank (or central bank), and a government. A Calvo lottery establishes which firms are allowed to re-optimize their prices. The outcome of the Calvo lotteries is assumed to be common knowledge among agents. The central bank supplies money to households so as to control the interest rate at which government's bonds pay out their return. Final goods producers buy intermediate goods from the firms, pack them into a final good to be sold to households and government in a perfectly competitive market. Households consume the final goods, demand money holdings from the central bank and bonds from the government, pay taxes to or receive transfers from the government, and supply labor to the firms.

There are aggregate and idiosyncratic shocks that hit the model economy. The aggregate shocks are: a technology shock, a monetary-policy shock, and a preference shock. The aggregate shocks have a persistent and white noise component.[6] All these shocks are orthogonal to the others at all leads and lags. Idiosyncratic shocks include the firm-specific technology, $A_{j,t}$, that determines the level of technology of firm j at time t, and the outcome of the Calvo lottery for price-optimization. The former shock is correlated with the aggregate technology shock. Both of the idiosyncratic shocks are orthogonal to each other at all leads and lags.

In section 3.1, I present the time protocol of the model. Section 3.2 presents the problem of the producers. Section 3.4 presents the price-setting problem of the intermediate-goods firms. Section 3.3 presents the problem of households. In Section 3.5, Central bank's behavior

[6] Monetary shocks are decomposed into an inflation-targeting shock and a white-noise shock to the Taylor rule. the latter is interpreted as a monetary policy shock.

is modeled. Section 3.6 deals with the log-linearization of the model equations. Section 3.7 discusses firm's signal-extraction problem within this log-lineariz(ed) set-up. Section C shows how to solve the model and analyzes the related issues. The case of perfect information is discussed in Section 3.10.

3.1 The Time Protocol

Any period t is divided into three stages. All actions that are taken in any given stage are simultaneous. At stage 0 $(t, 0)$, shocks realize and the central bank observes the realization of the aggregate shocks and sets the interest rate. At stage 1 $(t, 1)$, firms observe the realization of the idiosyncratic shocks (i.e., the firm-specific technology shock and the outcome of the Calvo lottery) and set their prices.[7] At stage 2 $(t, 2)$, households learn the realization of all the shocks in the economy and decide their consumption, $C_{i,t}$, money holdings, $M_{i,t}$, demand for government bonds,$B_{i,t}$ and labor supply, $N_{i,t}$. At this stage, firms learn the current level of inflation and real output, hire labor, and produce intermediate goods, $Y_{j,t}$ so as to deliver the demanded quantity of their good at the price they have set at the stage 1. Hence, intermediate-goods market, final-goods market, money market, bond market, and labor market clear.

3.2 Final-Goods Producers

The representative final-good producer combines a continuum of intermediate goods, $Y_{j,t}$ by using the technology:

$$Y_t = \left(\int_0^1 (Y_{j,t})^{\frac{\nu-1}{\nu}} \, dj \right)^{\frac{\nu}{\nu-1}} \tag{1}$$

where Y_t is the amount of the final good produced at time t, the parameter ν represents the elasticity of demand for each intermediate good and is assumed to be strictly larger than one. The producer takes the input prices, $P_{j,t}$, and output price, P_t, as given. Profit maximization implies that the demand for intermediate goods is:

$$Y_{j,t} = \left(\frac{P_{j,t}}{P_t} \right)^{-\nu} Y_t \tag{2}$$

where the competitive price of the final good, P_t, is given by

[7] Those firms that are not allowed to re-optimize their prices reset them according to the steady-state inflation rate.

$$P_t = \left(\int (P_{j,t})^{1-\nu} \, di \right)^{\frac{1}{1-\nu}}. \tag{3}$$

3.3 Households

Households solve:

$$\max_{C_{i,t}, B_{i,t}, M_{i,t}, N_{i,t}} \sum_{t=0}^{\infty} \beta^t \zeta_t \left[\ln C_{i,t} + \frac{\chi_m}{1-\gamma_m} \left(\frac{M_{i,t}}{P_t} \right)^{1-\gamma_m} - \chi_n N_{i,t} \right]$$

$$st$$

$$P_t C_{i,t} + B_{i,t} + M_{i,t} = W_t N_{i,t} + R_{t-1} B_{i,t-1} + M_{i,t-1} + \Pi_{i,t} - T_t$$

whereβ is the deterministic discount factor, ζ_t denotes a preference shock (or demand shock) that scales up or down the overall period utility, W_t is the (competitive) nominal wage, R_t stands for the interest rate, $\Pi_{i,t}$ are the dividends paid out by the firms, and T_t stands for government transfers. We can decompose the preference shock ζ_t into:

$$\zeta_t = g_t e^{\sigma_g \eta_{g,t}}$$

where $\eta_{g,t} \backsim \mathcal{N}(0,1)$ and

$$\ln g_t = \rho_g \ln g_{t-1} + \tilde{\sigma}_g \varepsilon_{g,t}$$

where $\varepsilon_{g,t} \backsim \mathcal{N}(0,1)$. We will denote $\hat{\zeta}_t = \ln \zeta_t$ and $\hat{g}_t = \ln g_t$. Since the dividends from firms, $\Pi_{i,t}$, are assumed to be equally shared among households, households turn out to face an identical problem. Thus, we can derive households' policy function as they were chosen by a representative agent.

The government transfers resources to/from and issue bonds to households. Furthermore, they decide their consumption in terms of final goods. Government spending is denoted by G_t. The government budget constraint is given by:

$$P_t G_t + R_{t-1} B_{t-1} - B_t + M_{t-1} - M_t = T_t$$

Since there is no capital accumulation and the Ricardo equivalence holds, the resource constraint implies $Y_t = C_t$.

3.4 Intermediate-Goods Firms' Price-Setting Problem

At the stage 1, intermediate goods firms observe the realization of their idiosyncratic technology shock and the outcome of the Calvo lottery,and set their prices. The firms commit themselves to satisfy any demanded quantity of their intermediate good that will arise in the stage 2 at the price they have set in the stage 1. Consider an arbitrary firm j. The real marginal costs for firm j are given by:

$$mc_{j,t} = \frac{W_t}{A_{j,t}P_t} \tag{4}$$

where $A_{j,t}$ is the technology shock that can be decomposed into (1) a trend component A_0, (2) a persistent aggregate component, z_t, (3) a white-noise aggregate component, $\eta_{a,t}$,and (4) a white-noise idiosyncratic component, $\eta^a_{j,t}$. More specifically, we have:

$$A_{j,t} = A_t e^{\eta^a_{j,t}} \tag{5}$$

with $A_0 > 1$, and $\eta^a_{j,t} \overset{iid}{\backsim} \mathcal{N}(0, \sigma^j_a)$, and

$$A_t = A^t_0 e^{z_t + \sigma_a \eta_{a,t}} \tag{6}$$

where $\eta_{a,t} \overset{iid}{\backsim} \mathcal{N}(0, 1)$ and

$$z_t = \rho_z z_{t-1} + \sigma_z \varepsilon_{z,t}$$

with $\varepsilon_{z,t} \overset{iid}{\backsim} \mathcal{N}(0, 1)$.

Firms face a Calvo lottery with probability θ of (sub-optimally) adjusting their price to the steady-state (gross) inflation rate, π_*. They set the price of the differentiated good they produce, given the linear technology:

$$Y_{j,t} = A_{j,t}N_{j,t} \tag{7}$$

where $N_{j,t}$ is the amount of labor employed by the firm j at time t.

Let me denote the firm j's nominal marginal costs at time t as $MC_{j,t} = W_t/A_{j,t}$, firm j's dividends as $\Pi_{j,t-1}$ and the outcome of the Calvo lottery for firm j at time t as $\mathbb{I}^{Calvo}_{j,t}$. Firm's information set $\mathcal{I}_{j,t}$ is defined below:

$$\mathcal{I}_{j,t} \equiv \{R_\tau, \pi_{\tau-1}, Y_{\tau-1}, W_{\tau-1}, \boldsymbol{\theta}_{j,\tau}, \Theta : \tau \leq t\} \tag{8}$$

where $\boldsymbol{\theta}_{j,t} \equiv \left\{A_{j,t}, P^*_{j,t}, Y_{j,t-1}, MC_{j,t-1}, N_{j,t-1}, \Pi_{j,t-1}, \mathbb{I}^{Calvo}_{j,t}\right\}$. Moreover, I denote, $\Xi_{t|t+s}$ as the time t value of one unit of the final good in period $t+s$ to the representative household.

Those firms that are allowed to re-optimize their price $P^*_{j,t}$ solve:

$$\max_{P^*_{j,t}} \mathbb{E}\left[\sum_{s=0}^{\infty} (\beta\theta)^s \,\Xi_{t|t+s} \left(\pi_*^s P^*_{j,t} - MC_{j,t+s}\right) Y_{j,t+s} | \mathcal{I}_{j,t}\right]$$

subject to the firm's specific demand in equation (2), to the production function (7).

3.5 The Central Bank

The central bank sets the interest rate R_t following a Taylor rule of type[8]:

$$R_t = R_{t-1}^{\rho_r} R_t^{*(1-\rho_r)} e^{\sigma_r \eta_{r,t}} \tag{9}$$

where $\eta_{r,t} \overset{iid}{\backsim} \mathcal{N}(0,1)$ and the desired nominal interest rate R^*_t is:

$$R^*_t = (r_* \pi^*_t) \left(\frac{\pi_t}{\pi^*_t}\right)^{\phi_\pi} \left(\frac{Y_t/Y_{t-1}}{A_0}\right)^{\phi_y} \tag{10}$$

where A_0 denotes the average growth rate of real output, π_t the inflation rate $\ln(P_t/P_{t-1})$, and π^*_t stands for the inflation target that is assumed to follow a stationary AR process:

$$\ln \pi^*_t = (1-\rho_\pi) \ln \pi_* + \rho_\pi \ln \pi^*_{t-1} + \sigma_\pi \varepsilon_{\pi,t}$$

where $\varepsilon_{\pi,t} \overset{iid}{\backsim} \mathcal{N}(0,1)$.

The central bank sets the interest rate R_t and is perfectly informed (i.e., it observes the contemporaneous realizations of all aggregate variables). We assume that the central bank cannot simply tell firms the realizations of the state variables (i.e., shocks, real output, and inflation) as there is an incentive for the central bank to lie to firms to generate surprise inflation with the aim of pushing output growth above the trend A_0.[9] Since there is not a commitment device that would back up central bank's words, then any central bank's statements about the real output, inflation, and shocks are not deemed as credible by the price-setting firms. If one assumes that there exists an exogenous commitment device (e.g., a law) that allows the central bank to commit herself to her statements, then the model would boil down to a (perfect-information) three-equations New Keynesian model. We will consider this perfect-commitment case as an interesting benchmark.

[8] We follow Arouba and Schorfheide (????) in specifying the monetary-policy reaction function.

[9] The fact that the central bank sets the interest rate before firms set their prices cannot be considered as a viable commitment device as the Taylor rule makes the interest rate to depend on output and inflation only up to a constant (i.e., the monetary policy shock $\eta_{r,t}$. Furthermore, the inflation targeting π^*_t is not observed by firms.

3.6 Detrending and log-linearization

Define the following stationary variables:

$$
\begin{aligned}
y_t &= \frac{Y_t}{A_0^t},\ c_t = \frac{C_t}{A_0^t};\ p_{j,t}^* = \frac{P_{j,t}^*}{P_t},\ y_{j,t} = \frac{Y_{j,t}}{A_0^t} \\
w_t &= \frac{W_t}{A_0^t P_t},\ a_t = \frac{A_t}{A_0^t}, R_t = \frac{R_t}{R_*},\ mc_{j,t} = \frac{MC_{j,t}}{P_t} \\
\xi_{j,t} &= A_0^t \Xi_{j,t}
\end{aligned}
$$

First, I solve firms and households' problems that are described in Sections 3.4 and 3.3 and obtain the consumption Euler equation and a price-setting equation. Second I detrend the non-stationary variables in these equations before log-linearize them around the perfect-information-deterministic steady-state equilibrium.

From the linearized price-setting equation, one can obtained an expression that resembles the New Keynesian Phillips curve, which is reported below (detailed derivations are in Appendix A).

$$
\hat{\pi}_t = (1-\theta)(1-\beta\theta)\sum_{k=0}^{\infty}(1-\theta)^k \widehat{mc}_{t|t}^{(k)} + \beta\theta\sum_{k=0}^{\infty}(1-\theta)^k \widehat{\pi}_{t+1|t}^{(k+1)} \tag{11}
$$

where $\widehat{\pi}_{t+1|t}^{(k)}$ denotes the average k-th order expectations about next period's inflation rate, $\widehat{\pi}_{t+1}$, that is $\widehat{\pi}_{t+1|t}^{(k)} \equiv \underbrace{\int \mathbb{E}_{j,t} \ldots \int \mathbb{E}_{j,t}}_{k} \widehat{\pi}_{t+1}$ and $\widehat{mc}_{t|t}^{(k)}$ denotes the average k-th order expectations about the real aggregate marginal costs, $\widehat{mc}_t$, which are defined as

$$
\widehat{mc}_{t|t}^{(k)} = \hat{y}_{t|t}^{(k+1)} - z_{t|t}^{(k)} - \sigma_a \eta_{a,t|t}^{(k)} \tag{12}
$$

It is easy to show that the log-linearized Euler equation is given by

$$
\hat{\zeta}_t - \widehat{y}_t = \mathbb{E}_t\hat{\zeta}_{t+1} - \mathbb{E}_t\widehat{y}_{t+1} - \mathbb{E}_t\hat{\pi}_{t+1} + \hat{R}_t \tag{13}
$$

The Taylor rule can be easily linearized:

$$
\hat{R}_t = \rho_r \hat{R}_{t-1} + (1-\rho_r)\left[\phi_\pi \hat{\pi}_t + (1-\phi_\pi)\hat{\pi}_t^* + \phi_y(\hat{y}_t - \hat{y}_{t-1})\right] + \sigma_r \eta_{r,t} \tag{14}
$$

where $\hat{\pi}_t^* \equiv \ln(\pi_t^*/\pi^*)$ and can be shown to follow $\hat{\pi}_t^* = \rho_\pi \hat{\pi}_{t-1}^* + \sigma_\pi \varepsilon_{\pi,t}$.

The parameter set of the model is given by the vector

$$\Theta = \left[\theta, \phi_\pi, \phi_y, \rho_r, \rho_z, \rho_g, \rho_\pi, \sigma_a, \sigma_z, \sigma_a^j, \sigma_a^j, \sigma_\pi, \sigma_r, \widetilde{\sigma}_g, \sigma_g\right]'$$

3.7 Firms' Signal Extraction Problem

The firms need to form beliefs about the current realization and the future dynamics of their nominal marginal costs given the observable signals in their information set $\mathcal{I}_{j,t}$. Characterizing how firms form such beliefs requires solving a signal extraction problem. It is important to notice that firms needs to form expectations on the price level to estimate nominal costs. Hence, they also need to form expectations what other price-setting firms expect about nominal marginal costs and on what other firms expect that other firms expect and so on (i.e., the so-termed higher-order expectations).

Let me denote the column vector $\varphi_{t|t}^{(0:k)}$ collecting the hierarchy of average higher-order expectations about the exogenous aggregate state variables (i.e., z_t, $\eta_{a,t}$, $\hat{\pi}^*_{t|t}$, $\eta_{r,t}$, g_t, $\eta_{g,t}$) up to the k-th order, that is,

$$\varphi_{t|t}^{(0:k)} \equiv \left[z_{t|t}^{(s)}, \eta_{a,t|t}^{(s)}, \hat{\pi}_{t|t}^{*(s)}, \eta_{r,t|t}^{(s)}, g_{t|t}^{(s)}, \eta_{g,t|t}^{(s)} : 0 \leq s \leq k\right]'$$

Solving firms signal extraction problem requires firms to have a perceived transition equation for the vector of aggregate state variables $\varphi_{t|t}^{(0:k)}$. We assume that firms are rational and they know the true transition equation of the state vector. We focus on equilibria where the vector of average higher-order expectations $\varphi_{t|t}^{(1:k)}$ follows a VAR(1) process, that is,

$$\varphi_{t|t}^{(0:k)} = \mathbf{M}\varphi_{t-1|t-1}^{(0:k)} + \mathbf{N}\boldsymbol{\varepsilon}_t \tag{15}$$

where $\boldsymbol{\varepsilon}_t$ is the vector of the i.i.d. shocks in the model, that is, $\boldsymbol{\varepsilon}_t \equiv \left[\begin{array}{cccccc} \varepsilon_{z,t} & \eta_{a,t} & \varepsilon_{\pi,t} & \eta_{r,t} & \varepsilon_{g,t} & \eta_{g,t} \end{array}\right]'$. Since the average zero-order expectations about a variable are conventionally defined as the variable itself, then, the transition equations driving the subvector $\varphi_t^{(0)} \equiv \left[z_t, \eta_{a,t}, \hat{\pi}^*_{t|t}, \eta_{r,t}, g_t, \eta_{g,t}\right]' \in \varphi_{t|t}^{(0:k)}$ is known. The remaining coefficients of the matrices $\mathbf{M}$ and $\mathbf{N}$ are yet to be determined.

Define $\mathbf{s}_t \equiv \left[\widehat{\pi}_t, \widehat{y}_t, \widehat{R}_t\right]'$ as the vector of inflation, output, and interest rate. Given the linear(ized) structure of the model obtained in Section 3.6, the endogenous state variables of the model (i.e., $\widehat{\pi}_t$, $\widehat{y}_t$, and $\widehat{R}_t$) can be shown to follow:

$$\hat{\pi}_t = \mathbf{a}_0\varphi_{t|t}^{(0:k)} + \mathbf{a}_1\mathbf{s}_{t-1} \tag{16}$$

$$\hat{y}_t = \mathbf{b}_0\varphi_{t|t}^{(0:k)} + \mathbf{b}_1\mathbf{s}_{t-1} \tag{17}$$

$$\hat{R}_t = \mathbf{c}_0\varphi_{t|t}^{(0:k)} + \mathbf{c}_1\mathbf{s}_{t-1} \tag{18}$$

where the vectors $\mathbf{a}_0$, $\mathbf{b}_0$, $\mathbf{c}_0$, $\mathbf{a}_1$, $\mathbf{b}_1$, $\mathbf{c}_1$are yet to be determined.

Given the transition equations (15)-(18) and firms' information set $\mathcal{I}_{j,t}$, firms' state-space model can be represented as follows:[10]

$$\mathbf{A}\underbrace{\begin{bmatrix} \varphi_{t|t}^{(0:k)} \\ \mathbf{s}_t \end{bmatrix}}_{\mathbf{X}_t} = \mathbf{B}\cdot\mathbf{X}_{t-1} + \mathbf{C}\cdot \underbrace{\begin{bmatrix} \varepsilon_{z,t} \\ \eta_{a,t} \\ \varepsilon_{\pi,t} \\ \eta_{r,t} \\ \varepsilon_{g,t} \\ \eta_{g,t} \end{bmatrix}}_{\boldsymbol{\varepsilon}_t} \tag{19}$$

$$\mathbf{A}=\begin{bmatrix} \mathbb{I} & \mathbf{0} \\ -\mathbf{v}_0 & \mathbb{I} \end{bmatrix},\ \mathbf{B}=\begin{bmatrix} \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{v}_1 \end{bmatrix},\ \ \mathbf{C}=\begin{bmatrix} \mathbf{N} \\ \mathbf{0} \end{bmatrix}$$

$$\begin{bmatrix} \ln A_{j,t} - \ln A_0 t \\ \pi_{t-1} - \ln \pi_* \\ \ln Y_{t-1} - \ln y_* - \ln A_0 t \\ \hat{R}_t - \ln \hat{R}_* \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_3^T \end{bmatrix}}_{\mathbf{D}_1}\mathbf{X}_t + \underbrace{\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_1^T \\ \mathbf{0} & \mathbf{1}_2^T \\ \mathbf{0} & \mathbf{0} \end{bmatrix}}_{\mathbf{D}_2}\mathbf{X}_{t-1} + \underbrace{\begin{bmatrix} \sigma_a^j \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{Q}}\eta_{j,t}^a \tag{20}$$

where $\mathbf{L} = \begin{bmatrix} 1, & 1, & \mathbf{0}_{1\times 6k+4} \end{bmatrix}$ and $\mathbf{v}_0 \equiv [\mathbf{a}_0', \mathbf{b}_0', \mathbf{c}_0']'$, $\mathbf{v}_1 \equiv [\mathbf{a}_1', \mathbf{b}_1', \mathbf{c}_1']'$.. Equation (19) collects the transition equations of the aggregate state vector of the model. Equation (20) is the observation equation that relates the variables that firms observe to the latent aggregate state variables. Note that only the first six rows of the matrices $\mathbf{M}$ and $\mathbf{N}$ are known as they are the transition equation of the six exogenous variables (i.e., $z_t, \eta_{a,t}, \hat{\pi}_{t|t}^*, \eta_{r,t}, g_t$, and $\eta_{g,t}$). Since the state-space model (19)-(20) is linear and all shocks are Gaussian, one can use the Kalman filter to solve firms' signal extraction problem and to obtain a law of motion of average higher-order expectations about the exogenous variables in the form of equation (15).

3.8 Model solution

Solving the model requires pinning down all the coefficients of the matrices $\mathbf{a}_0$, $\mathbf{b}_0$, $\mathbf{c}_0$, $\mathbf{a}_1$, $\mathbf{b}_1$, $\mathbf{c}_1$, $\mathbf{M}$ and $\mathbf{N}$ of the transition equations (15)-(18). The following assumption of common

[10]Note that we do not include the nominal wage W_{t-1} into the observables. In fact, this would be redundant because real output $\ln Y_t$ and inflation π_t are a sufficient statistic for the nominal wage W_t.

knowledge in rationality is necessary for solving the model (see also Nimark, 2008.)

Common Knowledge in Rationality It is true that $\underbrace{\int \mathbb{E}_{j,t} \ldots \int \mathbb{E}_{j,t}}_{s} \varphi^{(0:k)}_{t+h|t+h} = \mathbf{M}^h \varphi^{(s:k+s)}_{t|t}$.

This assumption ensures that firms use the actual transition equations of the average higher-order expectations to forecast the dynamics of the average expectations of higher orders.

The reduced-form transition equations (16)-(18) stem from the three structural equations of the model, that is, the IS equation (13), the Phillips curve (11), and the Taylor rule (14). As shown in Appendix B, the vectors $\mathbf{a}_0$, $\mathbf{b}_0$, $\mathbf{c}_0$, $\mathbf{a}_1$, $\mathbf{b}_1$, and $\mathbf{c}_1$ depend on the parameter values Θ and the matrices $\mathbf{M}$ and $\mathbf{N}$. More formally, we define that mapping ϕ, such that

$$[\mathbf{v}_0, \mathbf{v}_1]' = \phi\left(\mathbf{M}, \mathbf{N}; \Theta\right) \tag{21}$$

With this result in equation at hand, solving the model amounts to find a fixed point over the coefficients of matrices $\mathbf{M}$ and $\mathbf{N}$ as these are the only unknown matrices in the transition equations (15)-(18). To pin down the matrices $\mathbf{M}$ and $\mathbf{N}$, one has to solve firm's signal extraction problem which has been analyzed in Section 3.7. Appendix C shows in details that the application of the Kalman filter to solve firms' signal extraction problem and the assumption of common knowledge in rationality yield the following formula for the matrices $\mathbf{M}$ and $\mathbf{N}$:

$$\begin{aligned}
\mathbf{M} &= \begin{bmatrix} \mathbf{M}|_{(1:6,:)} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{6\times 6} & \mathbf{0}_{6\times 6k} \\ \mathbf{0}_{6k\times 6} & \left(\mathbf{W} - \mathbf{K}\mathbf{D}_1\mathbf{W} - \mathbf{K}\mathbf{D}_2\right)|_{(1:6k,1:6k)} \end{bmatrix} + \\
&\quad + \begin{bmatrix} \mathbf{0} \\ \mathbf{K}\left(\mathbf{D}_1\mathbf{W} + \mathbf{D}_2\right)|_{(1:6k,1:6(k+1))} \end{bmatrix} \qquad (22)\\
\mathbf{N} &= \begin{bmatrix} \mathbf{N}_{(1:6,:)} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{K}\mathbf{D}_1\mathbf{U}|_{(1:6k,1:6)} \end{bmatrix} \qquad (23)
\end{aligned}$$

where $\cdot|_{(n_1:n_2,m_1:m_2)}$ denotes the submatrix obtained by taking the elements lying between the n_1-th row and the n_2-th row and between the m_1-th column and the m_2-th column. The matrices $\mathbf{W} \equiv \mathbf{A}^{-1}\mathbf{B}$ and $\mathbf{U} = \mathbf{A}^{-1}\mathbf{C}$ denotes the reduced form dynamics of the model. The matrix $\mathbf{K}$ denotes the steady-state Kalman gain matrix, which is obtained from repeated applications of the Kalman filter to the state-space model (19)-(20).[11]

[11] Note that there are lagged state variables in firms' observation equations (20). Hence the Kalman-gain matrix K cannot be computed through the standard formula. A derivation of the formula has been provided

The algorithm used for solving the model is as follows: For given parameter values Θ:

1. Set $i = 1$ and guess the unknown coefficients of the matrices $\mathbf{M}$ and $\mathbf{N}$. Denote them as $\mathbf{M}^{(i)}$ and $\mathbf{N}^{(i)}$.

2. Use the mapping ϕ in equation (21) to pin down the matrices $[\mathbf{v}_0, \mathbf{v}_1]' = \phi\left(\mathbf{M}^{(i)}, \mathbf{N}^{(i)};\Theta\right)$.

3. Use the formulas (22)-(23) to obtain the new matrices $\mathbf{M}^{(i+1)}$ and $\mathbf{N}^{(i+1)}$

4. If max $\left(\left|vec\left(\mathbf{M}^{(i+1)}\right) - vec\left(\mathbf{M}^{(i)}\right)\right|, \left|vec\left(\mathbf{N}^{(i+1)}\right) - vec\left(\mathbf{N}^{(i)}\right)\right|\right) < \varepsilon$, where $\varepsilon > 0$ and small, then stop. Else set $i = i + 1$ and go to step 2.

The step 2 amounts to solve a system of non-linear equations on the coefficients of the matrices $\mathbf{v}_0$ and $\mathbf{v}_1$ (see Appendix B). In step 3, one needs to repeatedly apply the Kalman filter to firms' state-space model (19)-(20) to obtain the steady-state Kalman gain matrix $\mathbf{K}$ (see Appendix C).

3.9 The Channel Based on Affecting Firms' Expectations

The central bank can affect price setters' expectations by setting its policy rate is a salient feature of the model developed in the paper. This new transmission channel arises for the following three reasons. First, price setters have to form expectations about the evolution of their nominal marginal costs in order to set their prices because they face cost of price adjustment. Second, price setters perfectly observe the current policy rate and this is common knowledge among them. Third, the central bank follows a monetary policy rule with feedbacks.

Price setters use the policy rate as a signal both to figure out the current and future level[12] of the endogenous variables (i.e., inflation and real output), which affect the evolution of their nominal marginal costs, and to infer potential exogenous deviations from the rule (i.e., changes in the inflation target π_t^* and non-systematic deviations from the rule $\eta_{r,t}$). It is illustrative to re-write the Taylor rule (14) as follows:

$$\widetilde{R}_t = \underbrace{(1-\rho_r)\,\phi_\pi\hat{\pi}_t + (1-\rho_r)\,\phi_y\hat{y}_t}_{\textbf{Endogenous component of the signal}} + \underbrace{(1-\rho_r)\,(1-\phi_\pi)\,\hat{\pi}_t^* + \sigma_r\eta_{r,t}}_{\textbf{Exogenous component of the signal}} \tag{24}$$

where $\widetilde{R}_t \equiv \hat{R}_t - \rho_r\hat{R}_{t-1} + (1-\rho_r)\,\phi_y\hat{y}_{t-1}$ gathers all the variables in the rule that price setters know at time t. The first term on the right-hand side of the rule gathers all the endogenous variables, while the second term is affected only by exogenous variables.

by Nimark (2010). In the Appendix C I develop a Bayesian derivation of this formula.

[12] They care about future level of such a variables because they face cost of price adjustement.

When the variability endogenous component is larger than the that of the exogenous component, then a change in the policy rate will be interpreted by price setters as a response of the monetary authority to non-monetary shocks (i.e., technology and preference shocks) that have affected inflation and output. On the contrary, when most of the variability of the interest rate is explained by the exogenous component, then price setters will interpret observed changes in the policy rate as stemming from either modifications of the inflation target π_t^* or non-systematic deviations from the rule, $\eta_{r,t}$. These two interpretations have a drastically different impact on price setters' forecast about their nominal marginal costs and hence on their price-setting decisions. If price setters, for instance, believe that the policy rate has been increased because a strong positive demand shock has affected the economy, they will estimate rising nominal wages and growing marginal costs. On the contrary, if they interpret the rise in the policy rate as an exogenous deviations from the rule, then they will expect a rise in the real interest rate and hence a fall in their nominal marginal costs. Therefore, if the endogenous component of the policy signal is more volatile than the exogenous one, a monetary shock has delayed impacts on inflation and sluggish real effects because price setters learns about this shock only very slowly due to their initial misinterpretation.

The relative volatility of the endogenous component of the policy signal increases when the following occurs: (a) the central bank adopts a very proactive policy to either inflation or output (i.e., the values of the parameters ϕ_π or ϕ_y go up); (b) exogenous deviations from the rule are smaller in size on average (i.e., the variance parameters, σ_r and σ_π get smaller); (c) a change in non-policy parameters raising the unconditional variance of real output and inflation (e.g., the variance of technology or preference shocks gets larger).

Note that the direction of the effects (a) and (b) upon the importance of the endogenous component of the policy signal is ambiguous and varies with the values of other deep parameters. In fact, rising the degree of proactiveness of monetary policy or reducing the variability of monetary shocks will produce two opposite effects on the relative importance of the endogenous component. On one hand, they will increase the importance of the endogenous component by scaling up its variability or decreasing the variability of the exogenous component. On the other hand, very large values of these policy parameters will end up reducing the unconditional variance of real output or inflation.

It is finally important to notice that larger value for the coefficient on inflation, ϕ_π relative to that on output growth ϕ_y changes the information about the nature of the technology shocks that the policy rate conveys to price setters. Relatively large value for the parameter ϕ_π implies that the central bank will raise its policy rate after a negative technology shock as it cares mostly about inflation stability. When ϕ_π is relatively large, a rise in policy rate thus tends to coordinate agent expectations towards a negative technology shocks, which is expected to boost average expectations about real marginal costs and hence inflation. On the contrary, when ϕ_π is relatively small, then a rise of the interest rate tends to coordinate average expectations towards a positive technology shock that reduces inflation. A rise of the interest rate will always signal that a positive demand shock may have hit the economy. When price setters interpret the rise in the policy rate as coming form a demand shock, then expected real marginal costs and inflation will go up in the short run.

If price setters interpret a change in the policy rate as mainly coming from central bank's response to a demand shock via the endogenous component in equation (24), the effects of a monetary shock on inflation will be delayed. Effects on inflation tend to be amplified by the channel based on price setters' expectations, if price setters interpret a rise (fall) in the interest rate as coming from the response of the central bank to a positive (negative) technology shock.

3.10 The Case of Perfect Information

If the central bank could commit to her announcements, then she could reveal the state vector $\varphi_{t|t}^{(0:\infty)}$ to the firms. If such a commitment device existed, then price-setting firms would perfectly informed and the model would boil down into a three-equation new-Keynesian model with Calvo sticky-prices (add references). More specifically, the Phillips curve equation (11) would become $\hat{\pi}_t = \kappa_{pc}\widehat{mc}_t + \beta\mathbb{E}_t\widehat{\pi}_{t+1}$, where $\kappa_{pc} \equiv (1-\theta)(1-\theta\beta)/\theta$ and the real marginal costs $\widehat{mc}_t = \hat{y}_t - z_t - \sigma_a\eta_{a,t}$. The IS equation and the Taylor rule would be the same as in the incomplete information model. See equation (13) and (14). In this perfect information model the monetary shock propagates by affecting the intertemporal allocation of consumption and not by affecting beliefs. Real effects of money emerge as a result of price-stickiness as opposed to the sluggish adjustments of firms' expectations. In this model monetary shocks propagate through the channel based on the real interest rate, which is very typical to new Keynesian model. See Section 3.10. We call this standard New-Keynesian model as perfect information model or model with perfect commitment.

4 Empirical Analysis

This section contains the quantitative analysis of the model. I combine a prior distribution for the parameter set Θ with the likelihood function derived from the model and conduct Bayesian estimation of the parameters.

In Section 4.1, I present the data set and the state-space model for the econometrician.

In Section 4.2, I discuss the prior distribution for the model parameters. Section 4.3 presents the posterior distribution and the variance decomposition. In Section 4.4, I assess how well the model fits the observed inflation expectations. Section 4.5 studies the impulse response functions to monetary policy shocks. In Section 4.6, I discuss the effects of a structural policy of disinflation based on cutting the inflation target, π_t^*, pursued by the central bank.

4.1 Econometrician's State-Space model

Denote the matrices $\mathbf{W} \equiv \mathbf{A}^{-1}\mathbf{B}$ and $\mathbf{U} \equiv \mathbf{A}^{-1}\mathbf{C}$, where the characterization of the matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ has been discussed in Section C. The transition equations to be used for evaluating the likelihood function implied by the model are:

$$\left[\begin{array}{c} \varphi_{t|t}^{(0:k)} \\ \mathbf{s}_t \end{array}\right] = \mathbf{W} \cdot \left[\begin{array}{c} \varphi_{t-1|t-1}^{(0:k)} \\ \mathbf{s}_{t-1} \end{array}\right] + \mathbf{U} \cdot \boldsymbol{\varepsilon}_t$$

Two data sets are used for estimation. Both data sets include five observable variables: GDP growth rate, inflation, Federal Funds interest rate, one-quarter-ahead inflation expectations, and four-quarter-ahead inflation expectations. The last two observables are obtained from the *Survey of Professional Forecasters* (SPFs). A detailed description of the data set is provided in Table 1. The first data set $\mathcal{S}_1$ ranges from 1970:3 to 2008:4 while the second one $\mathcal{S}_2$ spans a period where the Federal Reserve has responded more aggressively towards inflation, that is 1983:2-2008:4. The paper wants to study how public's inflation expectations have reacted to unanticipated monetary shocks under different central bank's responsiveness to inflation deviations from the target. It has been widely documented that the Federal Reserve has been much more aggressive in fighting inflation during our second sample $\mathcal{S}_2$ than during the 1970s. Given that only few observations for 1970s are available, one cannot reliably estimate the model to a data set that ranges from 1970:3 to 1983:2. Given the larger variability of the observable variables during the 1970s, the posterior for model parameters when the full sample $\mathcal{S}_1$ is used, is likely to be highly affected the observations in the 1970s. The measurement equations are[13]:

$$100 \ln \frac{PGDP_t}{PGDP_{t-1}} = 100 \ln \pi_* + \hat{\pi}_t$$

$$\left[\ln\left(\frac{GDP_t}{POP_t^{\geq 16}}\right) - \ln\left(\frac{GDP_{t-1}}{POP_{t-1}^{\geq 16}}\right)\right] \cdot 100 = 100 \ln A_0 + \hat{y}_t - \hat{y}_{t-1}$$

[13]Note that the standard deviations of shocks and measurement errors are rescaled by a factor of 100.

$$100 \cdot FEDRATE_t = \hat{R}_t + 100 \ln R_*$$

$$\ln\left(\frac{PGDP3_t}{PGDP2_t}\right) 100 = \mathbf{1}_1^T \left[\mathbf{v}_0 \mathbf{T}^{(1)} \mathbf{M} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \left(\mathbf{v}_0 \mathbf{T}^{(1)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_{t-1}\right)\right] + 100 \ln \pi_* + \sigma_{m_1} \varepsilon_t^{m_1}$$

$$\ln\left(\frac{PGDP6_t}{PGDP2_t}\right) 25 = \mathbf{1}_1^T \left(\sum_{l=0}^{4} \mathbf{v}_1^{4-l} \mathbf{v}_0 \mathbf{T}^{(1)} \mathbf{M}^l \varphi_{t|t}^{(0:k)} + \mathbf{v}_1^5 \mathbf{s}_{t-1}\right) + 100 \ln \pi_* + \sigma_{m_2} \varepsilon_t^{m_2}$$

where $PGDP2_t$, $PGDP3_t$, $PGDP6_t$ are the *Survey of Professional Forecasters*[14] about the current, one-quarter-ahead, and four-quarter ahead GDP price index. We relate these statistics with the first moment of the distribution of firms' expectations implied by the model. $\varepsilon_t^{m_1}$ and $\varepsilon_t^{m_2}$ are two i.i.d. measurement errors, such that $\varepsilon_t^{m_1} \backsim \mathcal{N}(0,1)$ and $\varepsilon_t^{m_2} \backsim \mathcal{N}(0,1)$. These errors are intended to capture the difference between the observed expectations, which are the mean of the thirty professional forecasters' inflation expectations, and their model concepts, $\hat{\pi}_{t+1|t}^{(1)}$ and $\hat{\pi}_{t+4|t}^{(1)}$.

4.2 Priors

The prior medians and the 95% credible intervals are reported in Table 2. In the pre-sample from 1969:1 to 1970:2, I compute the average of real GDP and inflation to get a measure of $100 \ln A_0$ and $100 \ln \pi_*$. We centered the priors for the first two parameters at the value of these averages. Note that the discount factor β depends on the linear trend of real output A_0 and the steady-state real interest rate R_*/π_*. Hence, I fix the value for this parameter so as the steady-state nominal interest rate $\ln R_*$ matches the low-frequency behavior of $100 \cdot FEDRATE_t$ in the sample.

The prior median for the standard deviation of the idiosyncratic technology shocks, σ_a^j, is set so that the model, calibrated at the prior medians, matches the median absolute size of the U.S. finished goods producer price changes reported by Nakamura and Steinsson (2008). The priors for the standard deviations of the technology shocks, σ_z and σ_a, reflect the belief that most of the variance of the technology shocks is explained by the persistent component. This belief is consistent with what is the standard practice in the real business cycle literature (e.g., Kydland and Prescott, 1982).

The volatility of the monetary policy shock, σ_r, is very crucial as it affects not only the variability of the non-systematic component of monetary policy but also how much informative the interest rate R_t is regarding the current level of inflation, inflation target, and output. The prior for this parameter is centered at a smaller value than that for the volatility of the technology and preference shocks in order to reflect the belief that the

[14] Philly Fed Docs on SPFs to be cited here.

interest rate is informative for firms. The evolution of the inflation target, π_t^*, has to be interpreted as a long-run feature of the monetary policy strategy. Therefore, the prior for σ_π is set so that the inflation target π_t^* follows a process with very low variability. Priors for the remaining standard deviation of the i.i.d. shocks (i.e., $\widetilde{\sigma}_g$, σ_g, σ_{m1}, σ_{m2}) are informally taken according to the rule proposed by Del Negro and Schorfheide (2008) that the overall variance of endogenous variables is roughly close to that observed in the pre-sample.

The priors for the autoregressive parameters ρ_z, ρ_g, and ρ_π reflect the belief that the corresponding exogenous processes may exhibit sizeable persistence as it is usually observed in the macroeconomic data. Nonetheless, these priors are broad enough to accommodate a wide range of persistence degrees for these exogenous processes.

Priors for the parameters of the Taylor equation (i.e., response to inflation, ϕ_π, response to economic activity, ϕ_y, autoregressive parameter, ρ_r, and the standard deviation of the i.i.d. monetary shock, σ_r) are chosen as follows. The priors for ϕ_π and ϕ_y are centered at 2.00 and .25, respectively, and imply a fairly strong response to inflation and a moderate response to output growth. The prior for the autoregressive parameter, ρ_r is centered to 0.8, conjecturing that past monetary policy decisions have fairly persistent effects on current central bank's decision over the interest rate. The prior specification for the standard deviation of monetary shocks, σ_r, have been discussed above.

This prior distribution puts probability mass to values for the Calvo parameter θ implying that firms adjust their prices about every three quarters. This prior embodies information from micro studies on price-setting (Bils and Klenow, 2004, Klenow and Kryvtsov, 2008, Nakamura and Steinsson, 2008, and Klenow and Malin, 2010).

4.3 Posteriors

A closed-form expression for the posterior distribution is not available, but we can approximate the moments of the posterior distributions via the Metropolis-Hastings. $40,000$ posterior draws are obtained from the posterior. The posterior moments of parameters are reported in Table 3 when estimation is performed on the two samples $\mathcal{S}_1$and $\mathcal{S}_2$. Let me first describe the results for the full sample $\mathcal{S}_1$. Bayesian updating raises the volatility of the idiosyncratic component of technology shocks relative to that of the persistent component. The posterior median for the variance of the white-noise component of the technology shocks σ_a is larger than that for the idiosyncratic technology shock σ_a^j. These numbers imply a quite large signal-to-noise ratio relative to white-noise technology shocks $\eta_{a,t}$, suggesting that firms learn fairly quickly about these shocks. Nonetheless, firms find it harder to learn about the persistent component of technology z_t because the posterior median of σ_z is smaller relative to the posterior median of σ_a^j. Furthermore, the persistent preference shocks $\varepsilon_{g,t}$ is estimated to be much more volatile than the white-noise shock $\varepsilon_{g,t}$.

The posterior medians for the Taylor rule's parameters are by and large in line with what is often found by empirical studies that rely on perfect information models. Recall that the Taylor rule has two roles in the model. First, it describes central bank's reaction function. Second, it is an observation equation in firms' signal extraction problem. It is important to notice that the only source of information on preference shocks to firms is provided by the interest rate set by the central bank. Nonetheless, the posterior medians for the variance of the white-noise component of the preference shocks is relatively large. These parameter values tend to induce price-setters to believe that change in the interest rate has to be imputed to persistent preference shocks $\varepsilon_{g,t}$. The likelihood function perhaps selects these parameter values so as to allow the model to account for the sluggish adjustments of the observables variables, such as real GDP and inflation, to monetary policy shocks, which has been documented by Christiano, Eichenbaum, and Evans (2005). Furthermore, the Bayesian updating shrinks the response to inflation ϕ_π and raises the response to output ϕ_y, making the dynamics of the interest rate less informative about the current inflation and more informative about current output, relatively to what was conjectured in the prior.

The posterior medians for the parameters ρ_π and σ_π imply that the inflation target π_t^* follows a quite persistent and low-volatility process. These parameter values make sense in that the variable π_t^* intends to capture structural (i.e., very infrequent) changes of the monetary policy regime. The posterior median of the Calvo parameter implies that the average duration of price contracts is equal to about five quarters.

As far as the posterior distribution obtained from sample $\mathcal{S}_2$ is concerned, the following four facts are important to emphasize. First, Fed's response to inflation, ϕ_π, has increased in the subsample $\mathcal{S}_2$. Second, all shocks have smaller variance in the subsample $\mathcal{S}_2$. Third, in both samples the variance of the persistent preference shocks is larger than both that of the white-noise ones and that of monetary policy shock, implying that firms mainly interpret changes in the interest rate as central bank's responses to persistent preference shocks. Fourth, in both samples the variance of the white-noise technology shocks is larger than both that of the persistent shocks and that of the firm-specific shocks. This last result implies that firms find it harder to learn about persistent technology shocks relative to the white-noise ones.

Table 4 shows the variance decomposition of the observable variables (see Table 1) when the model is fitted to the sample $\mathcal{S}_1$.[15] Most of the variability of the growth rate of GDP is explained by the persistent preference shock $\eta_{g,t}$. The variability of the inflation rate,

[15]The variance decomposition is calculated by setting the parameter values to their posterior median.

the interest rate, one-quarter-ahead inflation expectations, and four-quarter-ahead inflation expectations are mainly explained by persistent technology shocks $\varepsilon_{z,t}$. The portions of variance of the observed inflation expectations that is explained by the two measurement errors $(\varepsilon_t^{m_1}, \varepsilon_t^{m_2})$ are rather small.

Table 5 shows the variance decomposition when the model is fitted to the subsample $\mathcal{S}_2$. Those shocks that explained most of the variability of the observables in the full sample $\mathcal{S}_1$ are still the dominant ones in the subsample too. Nonetheless, when one discards the observations in the 1970s, the technology shocks $(\varepsilon_{z,t}, \eta_{a,t})$ appear to contribute more to the variability of GDP growth and the persistent preference shocks $\varepsilon_{z,t}$ account for a larger portion of variability of inflation and inflation expectations.

4.4 Fitting the Observed Expectations

Table 6 reports the RMSE for the five observables when parameters are fixed at their posterior median (Full sample, $\mathcal{S}_1$). The model fits the SPFs better than the other three observables (i.e., GDP growth rate, inflation, and the Fed Funds rate). Figures 1 and 2 compare the observed expectations with the in-sample one-step ahead forecast of the model. It appears that the model forecasts the dynamics of inflation expectations rather well overall. Observed four-quarter ahead inflation expectations always lie within the 95% posterior credible set, while one-quarter-ahead inflation expectations are not captured by the model only for few quarters of the 1970s, especially when inflation expectations were rapidly rising. The model predict very well the evolution of inflation expectations in the last two decades. These findings empirically validate the model as a tool for studying the response of public expectations to policy shocks.

4.5 Propagation of Monetary Policy Shocks

Figure 3 shows the impulse response functions (and their 95% posterior credible sets in gray) of real GDP, inflation, interest rate, one-quarter-ahead inflation expectations, and four-quarter-ahead inflation expectations to a 25 basis-point rise in the interest rate. These impulse response functions are obtained when the model is fitted to the full sample (i.e., sample $\mathcal{S}_1$). The following three facts emerge: (a) the response of GDP and inflation to a monetary shock is very persistent; (b) the response of inflation and inflation expectations to a monetary shock is delayed and very sluggish, suggesting a great extent of price rigidity; (c) the initial response of inflation and inflation expectations to a contractionary monetary shock is positive. Fact (c) has been uncovered by Sims (1992) and dubbed as price puzzle by Eichenbaum (1992). In the subsequent sections we will discuss in details the explanation provided by the model for these three empirical facts that are common in the empirical literature based on VAR models.

4.5.1 Sluggish Adjustment of Output and Delayed Impact on Inflation

Facts (a) and (b) mainly stems from the new transmission channel based on inflation expectations. Figure 4 plots the 95% credible interval (the grey area reported in Figure 3) as well as the impulse response functions of the five observables to a monetary shock when price setters can observe the vector $\varphi_{t|t}^{(0:\infty)}$ errorless and this is common knowledge.[16] When this happens the incomplete information model boils down into a standard perfect-information New Keynesian model where nominal frictions are solely generated by price stickiness à la Calvo. In this model monetary shocks propagate through the channel based on the real interest rate, which is very typical to new Keynesian model. See Section 3.10. Figure 4 allows one to single out the transmission channel based on affecting price setters' expectations from the traditional one based on influencing the intertemporal allocation of consumption. It is apparent that the new channel that takes into account the responses of price setters' inflation expectations produces a great amount of persistence in the adjustment of GDP and delayed effects on inflation in the aftermath of a monetary shock.

Such a persistent pattern accords well with what is commonly found by the VAR literature[17]. It should be noted that this finding has been obtained by introducing information frictions into an otherwise very simple three-equation New Keynesian model. Within the class of New Keynesian DSGE models with perfect information, only large-scale DSGE models, such as Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007), with many shocks and frictions can reproduce sluggish and persistent response of macroeconomic variables to monetary disturbances. As shown in Figure, the baseline three-equation sticky-price model with perfect information fails generating such a pattern.

The introduction of the new channel based on price setters' expectations substantially raises the price rigidity and dampens the adjustment of real output upon impact. This raises the question about whether this price rigidity is too high. As a measure of price rigidity I compute by how many basis points (bps) the quarter-to-quarter inflation rate goes down at its peak after a monetary policy shock that reduces output by one percent (measured at the peak of the GDP response).[18] I call this statistic κ. The value for this statistic κ when one calibrates the model using the medians obtained when the incomplete information model is fitted to the full sample $\mathcal{S}_1$ (including the 1970s) is equal to 7.95 bps. This value for κ is

[16] The impulse response functions are computed at the posterior medians for the parameters of the incomplete information model estimated on a sample including observations over the 1970s.

[17] See, for instance, Christiano, Eichenbaum, and Evans (1999) and Stock and Watson (2001).

[18] This measure can be interpreted as the slope of the Phillips curve. See Schorfheide (2008).

line with what is found by the New Keynesian literature (i.e., 7-140 bps), as reported by Schorfheide (2008), albeit rather on the rigid side.

4.5.2 The Price Puzzle

To investigate the fact (c), Figure 5 plots the response of the vector collecting the average expectations of the exogenous variables up to the second order:[19]

$$\varphi_{t|t}^{(0:2)} \equiv \left[z_{t|t}^{(s)}, \eta_{a,t|t}^{(s)}, \hat{\pi}_{t|t}^{*(s)}, \eta_{r,t|t}^{(s)}, g_t^{(s)}, \eta_{g,t|t}^{(s)} : 0 \leq s \leq 2 \right]'$$

Average expectations about a persistent preference shocks strongly react after a monetary policy shock. This implies that price-setting firms are uncertain about interpreting the observed rise of the interest rate as a response of the central bank to a persistent preference shock that raises inflation and real output or as a monetary policy shock. Since a persistent preference shock would raise nominal marginal costs, firms will decrease their price less forcefully than what they would have done if they knew that a monetary shock has occurred (i.e., the case of perfect information). The lenient monetary policy against inflation conducted during the 1970s exacerbates the issue by increasing the inflationary effects of preference shocks. As a consequence, inflation rate does not react much to monetary shocks. Moreover, it is important to mention that firms hold strong beliefs that rise in the interest rate might be due to a negative transient technology shocks, $\eta_{a,t}$.

Furthermore, consider the following infinite-order moving-average representation for the endogenous state variables $\mathbf{s}_t \equiv \left[\hat{\pi}_t, \hat{y}_t, \hat{R}_t\right]'$ from the transition equations (16)-(18):

$$\mathbf{s}_t = \sum_{s=0}^{\infty} \mathbf{v}_1^s \mathbf{v}_0 \varphi_{t-s|t-s}^{(0:k)} \tag{25}$$

Given this decomposition, the exact (cumulative) contribution of the average higher-order expectations $\varphi_{t|t}^{(0:k)}$ (up to the order k) about the state of technology $\left(z_t, \eta_{a,t}\right)$, the state of monetary policy $\left(\pi_t^*, \eta_{r,t}\right)$, and the state of households' preferences $\left(g_t, \eta_{g,t}\right)$ to the responses of the endogenous variables, $\mathbf{s}_t \equiv \left[\hat{\pi}_t, \hat{y}_t, \hat{R}_t\right]'$, owing to a monetary policy shocks can be

[19] Recall that by convention the average zero-order expectations correspond to the realizations of the variables themselves, that is $\varphi_{t|t}^{(0)} = \varphi_t \equiv \left(z_t, \eta_{a,t}, \hat{\pi}_{t|t}^*, \eta_{r,t}, g_t, \eta_{g,t}\right)'$.

quantified as follows:

$$\begin{aligned}
\frac{\partial \mathbf{s}_{t+h}}{\partial \eta_{r,t}} \quad = \quad & \underbrace{\sum_{l=0}^{h}\left[\frac{\partial \mathbf{s}_{t+h}}{\partial z^{(0:k)}_{t+l|t+l}}\frac{\partial z^{(0:k)}_{t+l|t+l}}{\partial \eta_{r,t}}+\frac{\partial \mathbf{s}_{t+h}}{\partial \eta^{(0:k)}_{a,t+l|t+l}}\frac{\partial \eta^{(0:k)}_{a,t+l|t+l}}{\partial \eta_{r,t}}\right]}_{\equiv \text{ Component associated with the HOEs about the state of technology}} \\
& + \underbrace{\sum_{l=0}^{h}\left[\frac{\partial \mathbf{s}_{t+h}}{\partial \pi^{*(0:k)}_{t+l|t+l}}\frac{\partial \pi^{*(0:k)}_{t+l|t+l}}{\partial \eta_{r,t}}+\frac{\partial \mathbf{s}_{t+h}}{\partial \eta^{(0:k)}_{r,t+l|t+l}}\frac{\partial \eta^{(0:k)}_{r,t+l|t+l}}{\partial \eta_{r,t}}\right]}_{\equiv \text{ Component associated with the HOEs about the state of monetary policy}} \\
& + \underbrace{\sum_{l=0}^{h}\left[\frac{\partial \mathbf{s}_{t+h}}{\partial g^{(0:k)}_{t+l|t+l}}\frac{\partial g^{(0:k)}_{t+l|t+l}}{\partial \eta_{r,t}}+\frac{\partial \mathbf{s}_{t+h}}{\partial \eta^{(0:k)}_{g,t+l|t+l}}\frac{\partial \eta^{(0:k)}_{g,t+l|t+l}}{\partial \eta_{r,t}}\right]}_{\equiv \text{ Component associated with the HOEs about the state of preferences}} \qquad (26)
\end{aligned}$$

for $h = 0, 1, ..., 20$. Clearly, under perfect information the component associated with the Higher-Order Expectations (HOEs) about the state of technology and the component associated with the HOEs regarding the state of preferences are equal to zero in the aftermath of a monetary shock. For instance, under incomplete information, a large component associated with the HOEs about the state of preferences can be interpreted as a situation where price-setters mistakenly believe that the interest rate has changed as a result of a preference shock, which has inflationary effects (perhaps because of a monetary policy not tough enough against inflation).

Figure 6 shows why after a monetary tightening inflation does not fall as much as it would do under perfect information: firms mistakenly believe that preference shocks have occurred and have prompted the central bank to raise the interest rate. This can be seen by observing that the response of the component associated with the HOEs about the state of preferences is positive and quite strong, especially upon shock. This pattern partially offsets the response of the component associated with the HOEs about the state of monetary policy, which pushes the response of inflation into the negative region. The response of the component associated with the HOEs about the state of technology to a contractionary monetary policy is relatively small, suggesting that inflation rate does not tend to rise because firms might mistakenly perceive the rise of the interest rate as a consequence of a negative technology shock. Furthermore, the graph on top shows that the same reason can be applied to explain why real output only moderately falls after a monetary policy shock.

4.5.3 The Disappearance of the Price Puzzle after the 1970s

Recent studies based on VAR models[20] have documented that the price puzzle is statistically relevant in the 1970s only. I compute the impulse response functions to a contractionary monetary policy shock when I estimate the model not including the 1970s (i.e., data set $\mathcal{S}_2$). Posterior medians for this subsample are reported in Table 3 and were discussed in Section 4.3. Figure 7 reports the impulse response functions (and their 95% posterior credible sets) to a 25 bps rise of the policy rate after 1970s. The main finding is that excluding the data of the 1970s from the sample leads the (average first-order) inflation expectations to respond negatively to monetary shocks upon impact. From 1980s on, the Federal Reserve, hence, seems to have successfully affirmed its control over the initial response of price setters' inflation expectations. One-step ahead inflation expectations never lie in the positive region (i.e., values for inflation expectations that are larger than their pre-shock level). Nonetheless, the response of four-quarter-ahead inflation expectations to a monetary policy shock shifts down only by a little when compared to their response in the full sample that includes the 1970s.

This finding is line with other VAR studies but raises the question of what can explain the disappearance of the positive initial response of inflation to monetary shock? A quick comparison of Figure 8 with Figure 6 shows that the average higher order expectations about the monetary component react more strongly upon the monetary than what it was when one includes the 1970s in the data range. Furthermore, after 1970s price setters consider, on average, less likely that the rise in the policy rate has been triggered by a positive and persistent preference shocks. These two effects causes inflation to respond negatively upon a monetary shock.

4.6 Structural Monetary Policies of Disinflation

In this section we study the macroeconomic effects of suddenly introducing a new monetary policy regime which is characterized by a fall in the inflation target $\widehat{\pi}_t^*$ by 50 bps. This policy experiment is performed by re-scaling the shock ε_t^π so that π_t^* falls by 50 bps. We perform this experiment in three scenarios: the case of incomplete information when the

[20] Price puzzle is found to be statistically insignificant after 1970s by several empirical studies based on VAR models. See for instance, Barth and Ramey (2001), Hansen (2004), Castelnuovo and Surico (2010). and Ravn, Schmitt-Grohe, Uribe, and Uusküla (2010). Barth and Ramey (2001) and Hansen (2004) document the disappearence of the price puzzle by using monthly data and zero-restrictions to identify the monetary shock. Hansen (2004) corroborates this result by using several proxies for the expected future inflation. Castelnuovo and Surico (2010) use sign restrictions as proposed by Uhlig (2005) to identify the shock. Ravn, Schmitt-Grohe, Uribe, and Uusküla (2010) find that afetr 1970s the price puzzle is substantially weaker. These scholars argue that the mitigation of the price puzzle observed after the 1970s is due to structural changes in the conduct of monetary policy.

model is estimated to the full sample $\mathcal{S}_1$, the case of incomplete information when the model is estimated to the sub-sample $\mathcal{S}_2$, and the case of perfect information (using the parameter values of the incomplete information model fitted to the full sample $\mathcal{S}_1$). The first scenario is interpreted as a description of a policy of disinflation to be conducted during the 1970s under no commitment. The second scenario reproduces the effect of a structural policy of disinflation after 1970s under no commitment. The third scenario can be interpreted as the case when the central bank is endowed with a credible commitment device associated with the disinflation policy (e.g., the inflation target is lowered by an enforceable law).

The responses of real GDP, inflation, interest rate, one-quarter-ahead inflation expectations, and four-quarter-ahead inflation expectations to a long-run disinflation policy that changes the inflation target π_t^* by a half of a percentage point are reported in Figure 9. In this figure we use the posterior distribution based on the full sample (i.e., $\mathcal{S}_1$). The lack of commitment makes the costs of disinflation in terms of output loss much more prolonged. Such costs clearly emerge in the medium and long run (i.e., from three quarters after the shock onward). Under perfect commitment most of the output loss occurs within the first two years. Disinflating the economy in the 1980s under no commitment is more costly than doing this in the 1970s with or without[21] a perfect commitment: The output loss is 10 times larger than that in case of perfect commitment. The lack of commitment makes the disinflation policy unsuccessful: after five years inflation has fallen only by 6 bps (17 bps if the policy is implemented after the 1970s), which is well below the disinflation target of 50 bps. Under perfect commitment the inflation rate overshoots the inflation target of 50 bps suddenly after the new regime is introduced. The disinflation policy under perfect commitment attains its goal in about one year and half. Under no commitment the central bank has to raise the interest rate to disinflate. In contrast, under perfect commitment the simple announcement of the disinflation policy strongly curbs inflation (see the responses of inflation expectations under perfect commitment) without the need for a rise of the interest rate, which is actually cut. Finally, in absence of an effective commitment device, price setters' inflation expectations adjust very sluggishly to the new inflation target.

To investigate the reason why the model predicts that a policy of disinflation is doomed in absence of a perfect commitment, we take a look at how higher-order expectations about the exogenous variables react to a disinflation policy when the central bank cannot commit. Figure 10 shows the impulse response functions of the average expectations from order zero up to the second order, that is $\varphi_{t|t}^{(0:2)}$ when one considers only the scenario 1 (i.e., full sample and no commitment). We observe two facts. First, the disinflation policy dramatically fails influencing firms' expectations about the inflation target, π_t^* under lack of commitment.

[21] This result is largely due to a more aggressive monetary policy to inflation deviations from its trend π_t^*.

Second, it is clear that price-setting firms interpret the rise in the interest rate as a central bank's response to persistent preference shocks (see the right-bottom graph). Nonetheless, this second effect is dominated by the first one as an inspection of Figures 11 and 12 reveals. The component associated with the HOEs about the state of preferences only marginally curbs the response of output and inflation to the policy regime change. The dominant effect is the lack of adjustment of the HOE component associated with the state of monetary policy. This can be seen by noting the large distance in between this component (i.e., the line with white circles) and the inflation target (i.e., the horizontal line with crosses).

Figure 13 shows the responses of the five observables to an inflation target shock under several values for the coefficient on inflation in the Taylor rule. Such a counterfactual experiment delivers that stronger responses to inflation cause price setters to learn faster about the new level of the inflation target. One can indeed see that average inflation expectations are more reactive when the central bank reacts more forcefully against inflation. Nonetheless, faster learning implies a larger cost in terms of output loss when the central bank is very aggressive against inflation.

To sum up, the failure of the disinflation policy (i.e., high costs in terms of output loss and small reduction of inflation) under no commitment is due to the sluggish reaction of the price-setters' (average) expectations about the inflation target. The reason for such a sluggish adjustment is that price-setting firms consider a change in the inflation target π_t^* as an extremely rare event. This is captured by the very small standard deviation σ_π (see Table 3). Since firms face a signal extraction problem, they tend to learn very slowly any changes affecting the long-run component of monetary policy. It clearly emerges that a powerful remedy to reduce the output loss and increase the effectiveness of disinflation policies is to construct a viable commitment device to back up such policies. When the central bank is very aggressive to inflation, structural policies of disinflation are effective to a greater extent but the cost in terms of output loss is substantial.

5 Concluding Remarks

TBW

References

Adam, K. (2007): "Optimal Monetary Policy with Imperfect Common Knowledge," *Journal of Monetary Economics*, 54(2), 267–301.

——— (2009): "Monetary Policy and Aggregate Volatility," *Journal of Monetary Economics*, 56(S1), S1–S18.

Amato, J. D., and H. S. Shin (2003): "Public and Private Information In Monetary Policy Models," BIS Working Paper series 138.

——— (2006): "Imperfect Common Knowledge and the Information Value of Prices," *Economic Theory*, 27, 213–241.

Angeletos, G.-M., and J. La'O (2009): "Incomplete Information, Higher-Order Beliefs, and Price Inertia," *Journal of Monetary Economics*, 56(S1), S19–S37.

Barth, M., and V. Ramey (2001): "The Cost Channel of Monetary Transmission," *NBER Macroeconomics Annual*, 16, 199–240.

Bernanke, B. S. (2004): "The Great Moderation," Washington, DC: remarks delivered at the Meeting of the Eastern Economic Association.

Bils, M., and P. J. Klenow (2004): "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, 112(5), 947–985.

Calvo, G. A. (1983): "Staggered Prices in a Utility Maximizing Framework," *Journal of Monetary Economics*, 12(3), 383–398.

Castelnuovo, E., and P. Surico (2010): "Monetary policy, inflation expectations and the price puzzle," *Economic Journal*, 120, 1262Ű1283.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (1999): "Monetary Policy Shocks: What Have We Learned and To What End?," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1A, pp. 65–148. North Holland, New York.

——— (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113(1), 1–45.

Del Negro, M., and F. Schorfheide (2008): "Forming priors for DSGE models (and how it affects the assessment of nominal rigidities)," *Journal of Monetary Economics*, 55(7), 1191–1208.

Eichenbaum, M. (1992): "Comment on Interpreting the macroeconomic time series facts: the effects of monetary policy," *European Economic Review*, 36(5), 1001–1011.

Gorodnichenko, Y. (2008): "Endogenous Information, Menu Costs, and Inflation Persistence," NBER Working paper 14184.

Hansen, G. D. (2004): "The price puzzle reconsidered," *Journal of Monetary Economics*, 51, 1385–1413.

Hellwig, C. (2002): "Public Announcements, Adjustment Delays and the Business Cycle," UCLA Mimeo.

Judd, K. L. (1998): *Numerical Methods in Economics*. The MIT Press, Boston.

King, M. (2005): "Monetary Policy: Practice Ahead of Theory," MAIS Lecture.

Klenow, P. J., and O. Kryvtsov (2008): "State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?," *Quarterly Journal of Economics*, 123(3), 863–904.

Klenow, P. J., and B. A. Malin (2010): "Microeconomic Evidence on Price-Setting," in *Handbook of Monetary Economics*, ed. by B. Friedman, and M. Woodford, vol. 3A. North Holland, New York.

Kydland, F. E., and E. C. Prescott (1982): "Time To Build and Aggregate Fluctuations," *Econometrica*, 50(6), 1345–1370.

Lorenzoni, G. (2009): "A Theory of Demand Shock," *American Economic Review*, 99(5), 2050–2084.

——— (2010): "Optimal Monetary Policy with Uncertain Fundamentals and Dispersed Information," *Review of Economic Studies*, 77(1), 305–338.

Lucas, R. E. J. (1972): "Expectations and the Neutrality of Money," *Journal of Economic Theory*, 4(2), 103–124.

Maćkowiak, B., E. Moench, and M. Wiederholt (2009): "Sectoral Price Data and Models of Price Settings," *Journal of Monetary Economics*, 56(S)(S1), 78–99.

Maćkowiak, B., and M. Wiederholt (2009): "Optimal Sticky Prices under Rational Inattention," *American Economic Review*, 99(3), 769–803.

——— (2010): "Business Cycle Dynamics under Rational Inattention," CEPR Discussion Paper No. 7691.

Mankiw, G. N., and R. Reis (2002a): "Sticky Information: A Model of Monetary Nonneutrality and Structural Slumps," in *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, ed. by P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, pp. 64–86. Princeton University Press, Princeton.

——— (2002b): "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics*, 117(4), 1295–1328.

——— (2006): "Pervasive Stickiness," *American Economic Review*, 96(2), 164–169.

——— (2007): "Sticky Information in General Equilibrium," *Journal of the European Economic Association*, 5(2-3), 603–613.

Melosi, L. (2010): "Estimated Models with Information Frictions," mimeo.

Morris, S., and H. S. Shin (2003a): "Global Games: Theory and Applications," in *Advances in Economics and Econometrics, the Eighth World Congress*, ed. by M. D. et al. Cambridge University Press, Cambridge.

——— (2003b): "Social Value of Public Information," *American Economic Review*, 92(5), 1521–1534.

——— (2008): "Coordinating Expectations in Monetary Policy," in *Central Banks as Economic Institutions*, ed. by J.P.Touffut. Edward Elgar Publishing, Northampton.

Nakamura, E., and J. Steinsson (2008): "Five Facts about Prices: A Reevaluation of Menu Cost Models," *Quarterly Journal of Economics*, 123(4), 1415–1464.

Nimark, K. (2008): "Dynamic Pricing and Imperfect Common Knowledge," *Journal of Monetary Economics*, 55(8), 365–382.

——— (2010): "A Low Dimensional Kalman Filter For Systems with Lagged Observables," Mimeo.

Paciello, L. (2010): "Monetary Policy and Price Responsiveness to Aggregate Shocks under Rational Inattention," EIEF Mimeo.

Phelps, E. S. (1970): *Microeconomic Foundations of Employment and Inflation Theory*. Macmillan, London.

Ravn, M., S. Schmitt-Grohe, M. Uribe, and L. Uusküla (2010): "Deep habits and the dynamic effects of monetary policy shocks," *Journal of the Japanese and International Economies*, 24, 236–258.

Reis, R. (2006a): "Inattentive Consumers," *Journal of Monetary Economics*, 53(8), 1761–1800.

——— (2006b): "Inattentive Producers," *Review of Economic Studies*, 73(3), 793–821.

——— (2009): "A Sticky-Information General-Equilibrium Model for Policy Analysis," NBER Working paper 14732.

Rondina, G. (2008): "Incomplete Information and Informative Pricing," University of California San Diego, Mimeo.

Schorfheide, F. (2008): "DSGE Model-Based Estimation of the New Keynesian Phillips Curve," *Economic Quarterly*, 94(4), 397–433.

Sims, C. A. (1992): "Interpreting the macroeconomic time series facts: the effects of monetary policy," *European Economic Review*, 36(5), 975–1000.

Smets, F., and R. Wouters (2007): "Shocks and Frictions in US Business Cycles: A Bayeasian DSGE Approach," *American Economic Review*, 97(3), 586–606.

Stock, J. H., and M. W. Watson (2001): "Vector Autoregressions," *Journal of Economic Perspectives*, 15(4), 101–115.

Townsend, R. M. (1983a): "Equilibrium Theory with Learning and Disparate Information," in *Individual Forecasting and Aggregate Outcomes*, ed. by R. Frydman, and E. S. Phelps, pp. 169–197. Cambridge University Press, Cambridge.

——— (1983b): "Forecasting the Forecasts of Others," *Journal of Political Economy*, 91(4), 546–588.

Trichet, J.-C. (2006): "The role of money Ű money and monetary policy in the 21st century," Speech at the panel intervention at the Fourth ECB Central Banking Conference ŞThe role of money: money and monetary policy in the twenty-first centuryŤ.

Uhlig, H. (2005): "What are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure," *Journal of Monetary Economics*, 52, 381–419.

WOODFORD, M. (2002): "Imperfect Common Knowledge and the Effects of Monetary Policy," in *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, ed. by P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, pp. 25–58. Princeton University Press, Princeton.

——— (2005): "Central-Bank Communication and Policy Effectiveness," Federal Reserve Bank of Kansas City Symposium mimeo.

——— (2009): "Information-Constrained State-Dependent Pricing," *Journal of Monetary Economics*, 56(S)(S1), S100–S124.

Tables and Figures

Table 1: Observables

Variables	Description	Source
GDP_t	Gross Domestic Product - Quarterly	BEA (GDPC96)
$POP_t^{\geq 16}$	Civilian nonistitutional population - 16 years and over	BLS (LNS10000000)
$PGDP_t$	Consumer Price Index - Averages of Monthly Figures	BLS (CPIAUCSL)
$FEDRATE_t$	Effective Federal Funds Rate - Averages of Daily Figures	Board of Governors (FEDFUNDS)
$PGDP2_t$	Mean of Expections of current GDP price index	SPFs in mean.xls (PGDP2)
$PGDP3_t$	Mean of Expections of one-quarter-ahead GDP price index	SPFs in mean.xls (PGDP3)
$PGDP6_t$	Mean of Expectations of one-year-ahead GDP price index	SPFs in mean.xls (PGDP6)

Table 2: Prior Distributions

Name	Support	Density	Median	95% Interval
θ	$[0,1]$	Beta	0.65	$[0.55, 0.93]$
ϕ_π	$\mathbb{R}+$	Gamma	1.80	$[0.89, 2.80]$
ϕ_y	$\mathbb{R}+$	Gamma	0.25	$[0.07, 0.45]$
ρ_r	$[0,1]$	Beta	0.80	$[0.60, 0.97]$
ρ_z	$[0,1]$	Beta	0.85	$[0.65, 0.99]$
ρ_g	$[0,1]$	Beta	0.85	$[0.65, 0.99]$
ρ_π	$[0,1]$	Beta	0.85	$[0.65, 0.99]$
σ_a	$\mathbb{R}+$	InvGamma	0.10	$[0.05, 0.24]$
σ_z	$\mathbb{R}+$	InvGamma	0.70	$[0.51, 0.97]$
σ_a^j	$\mathbb{R}+$	InvGamma	2.50	$[1.00, 4.85]$
σ_π	$\mathbb{R}+$	InvGamma	0.02	$[0.01, 0.03]$
σ_r	$\mathbb{R}+$	InvGamma	0.10	$[0.05, 0.25]$
σ_g	$\mathbb{R}+$	InvGamma	0.25	$[0.10, 0.48]$
$\tilde{\sigma}_g$	$\mathbb{R}+$	InvGamma	0.25	$[0.10, 0.48]$
σ_{m_1}	$\mathbb{R}+$	InvGamma	0.25	$[0.10, 0.48]$
σ_{m_2}	$\mathbb{R}+$	InvGamma	0.25	$[0.10, 0.48]$
$\ln A_0$	$\mathbb{R}$	Normal	0.00	$[-0.20, 0.20]$
$\ln \pi_*$	$\mathbb{R}$	Normal	0.00	$[-0.20, 0.20]$

Table 3: Posterior Distributions

Name	Full Sample ($\mathcal{S}_1$)		Sub-sample ($\mathcal{S}_2$)	
	Median	95% Interval	Median	95% Interval
θ	0.82	$[0.79, 0.84]$	0.85	$[0.83, 0.86]$
ϕ_π	1.42	$[1.40, 1.46]$	1.81	$[1.64, 2.08]$
ϕ_y	0.48	$[0.40, 0.56]$	0.67	$[0.56, 0.75]$
ρ_r	0.74	$[0.70, 0.76]$	0.83	$[0.80, 0.85]$
ρ_z	0.99	$[0.98, 0.99]$	0.99	$[0.99, 1.00]$
ρ_g	0.77	$[0.75, 0.79]$	0.79	$[0.76, 0.81]$
ρ_π	0.99	$[0.99, 1.00]$	0.99	$[0.99, 1.00]$
σ_a	5.98	$[4.16, 7.85]$	5.65	$[5.05, 6.60]$
σ_z	1.29	$[1.05, 1.54]$	0.98	$[0.80, 1.16]$
σ_a^j	4.33	$[3.36, 5.41]$	3.07	$[2.71, 3.38]$
σ_π	0.04	$[0.03, 0.04]$	0.02	$[0.01, 0.03]$
σ_r	0.25	$[0.22, 0.28]$	0.13	$[0.11, 0.15]$
σ_g	0.18	$[0.10, 0.28]$	0.18	$[0.10, 0.24]$
$\tilde{\sigma}_g$	3.70	$[3.20, 4.15]$	2.46	$[2.21, 2.73]$
σ_{m_1}	0.14	$[0.13, 0.16]$	0.13	$[0.12, 0.14]$
σ_{m_2}	0.13	$[0.12, 0.15]$	0.12	$[0.11, 0.13]$
$100 \ln A_0$	0.28	$[0.25, 0.31]$	0.34	$[0.27, 0.42]$
$100 \ln \pi_*$	0.37	$[0.03, 0.74]$	1.09	$[0.62, 1.61]$

Table 4: Variance Decomposition (contribution in percentage): Full Sample

Observables	Shocks							
	$\varepsilon_{z,t}$	$\eta_{a,t}$	$\varepsilon_{\pi,t}$	$\eta_{r,t}$	$\varepsilon_{g,t}$	$\eta_{g,t}$	$\varepsilon_t^{m_1}$	$\varepsilon_t^{m_2}$
GDP growth	4.90	7.85	0.01	4.70	77.21	5.32	0.00	0.00
Inflation	70.39	14.64	4.26	1.15	9.52	0.05	0.00	0.00
Fed Funds Rate	74.21	1.94	1.21	12.71	9.86	0.08	0.00	0.00
Infl Exp's (1q ahead)	79.98	2.37	3.59	1.03	6.81	0.03	6.19	0.00
Infl Exp's (4q ahead)	82.62	2.39	2.93	1.37	5.55	0.02	0.00	5.12

Table 5: Variance Decomposition (contribution in percentage): Sub-sample

Observables	Shocks							
	$\varepsilon_{z,t}$	$\eta_{a,t}$	$\varepsilon_{\pi,t}$	$\eta_{r,t}$	$\varepsilon_{g,t}$	$\eta_{g,t}$	$\varepsilon_t^{m_1}$	$\varepsilon_t^{m_2}$
GDP growth	7.46	17.36	0.02	2.71	62.90	9.56	0.00	0.00
Inflation	79.67	11.64	2.86	1.37	4.42	0.05	0.00	0.00
Fed Funds Rate	89.66	0.54	0.41	4.50	4.82	0.08	0.00	0.00
Infl Exp's (1q ahead)	86.14	1.22	2.12	0.94	2.50	0.03	7.05	0.00
Infl Exp's (4q ahead)	88.84	0.85	1.53	0.69	2.61	0.02	0.00	5.45

Table 6: RMSE - Full Sample

Observables	RMSE
GDP growth	0.83
Inflation	0.30
Fed Funds Rate	0.26
One-Quarter Ahead Inflation Expectation (SPFs)	0.15
Four-Quarter Ahead Inflation Expectation (SPFs)	0.13

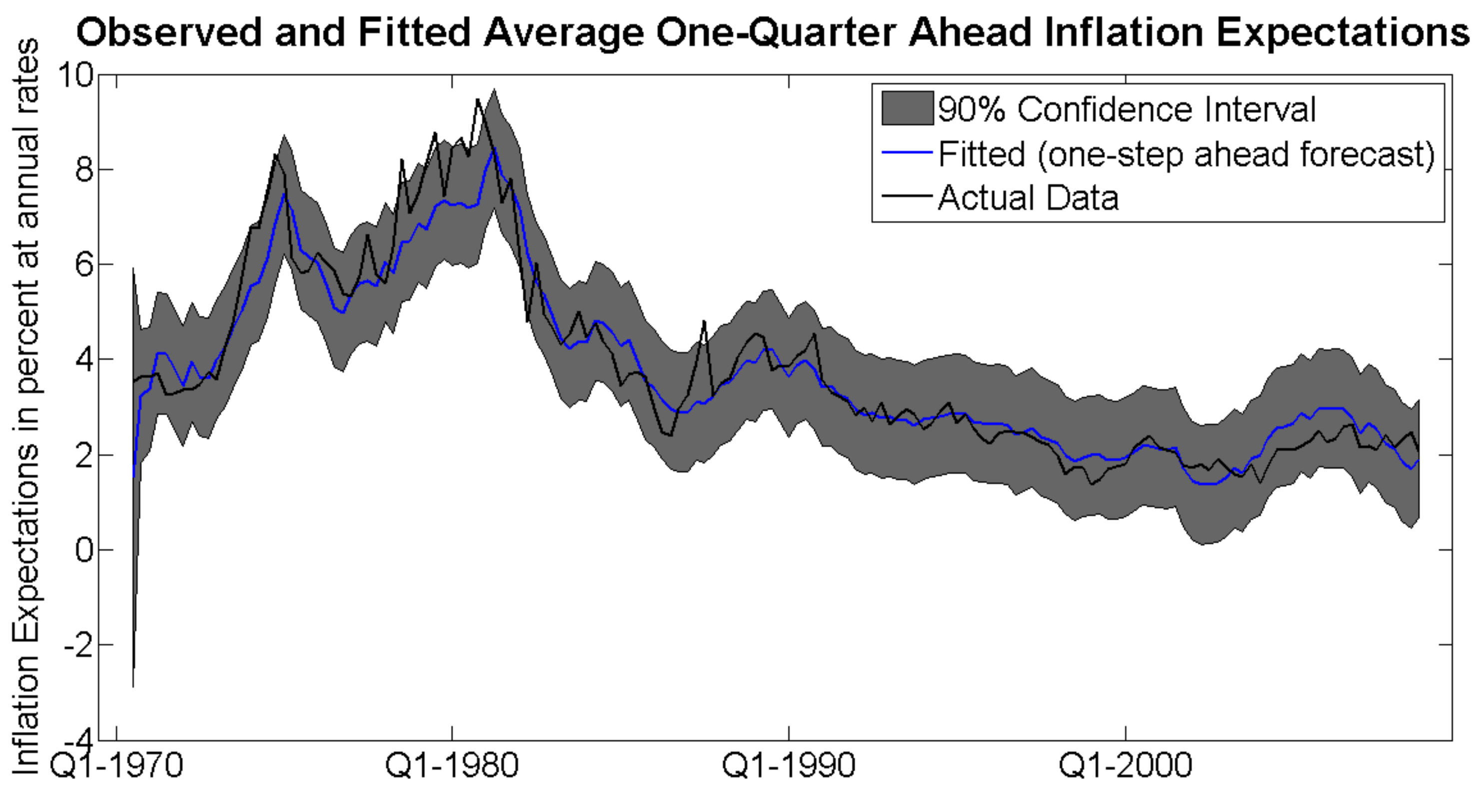

Figure 1: Observed and Fitted One-Quarter-Ahead Average Inflation Expectations

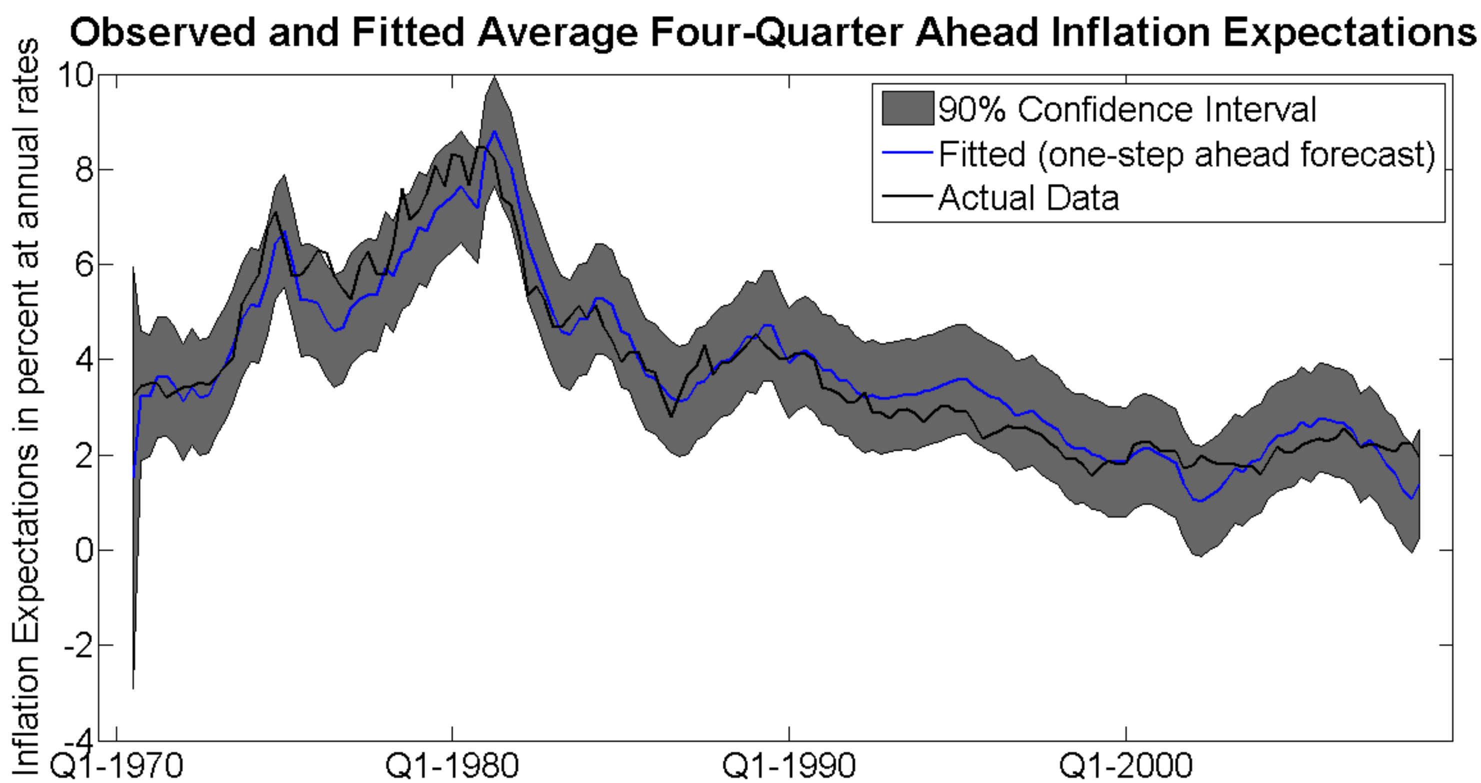

Figure 2: Observed and Fitted Four-Quarter-Ahead Average Inflation Expectations

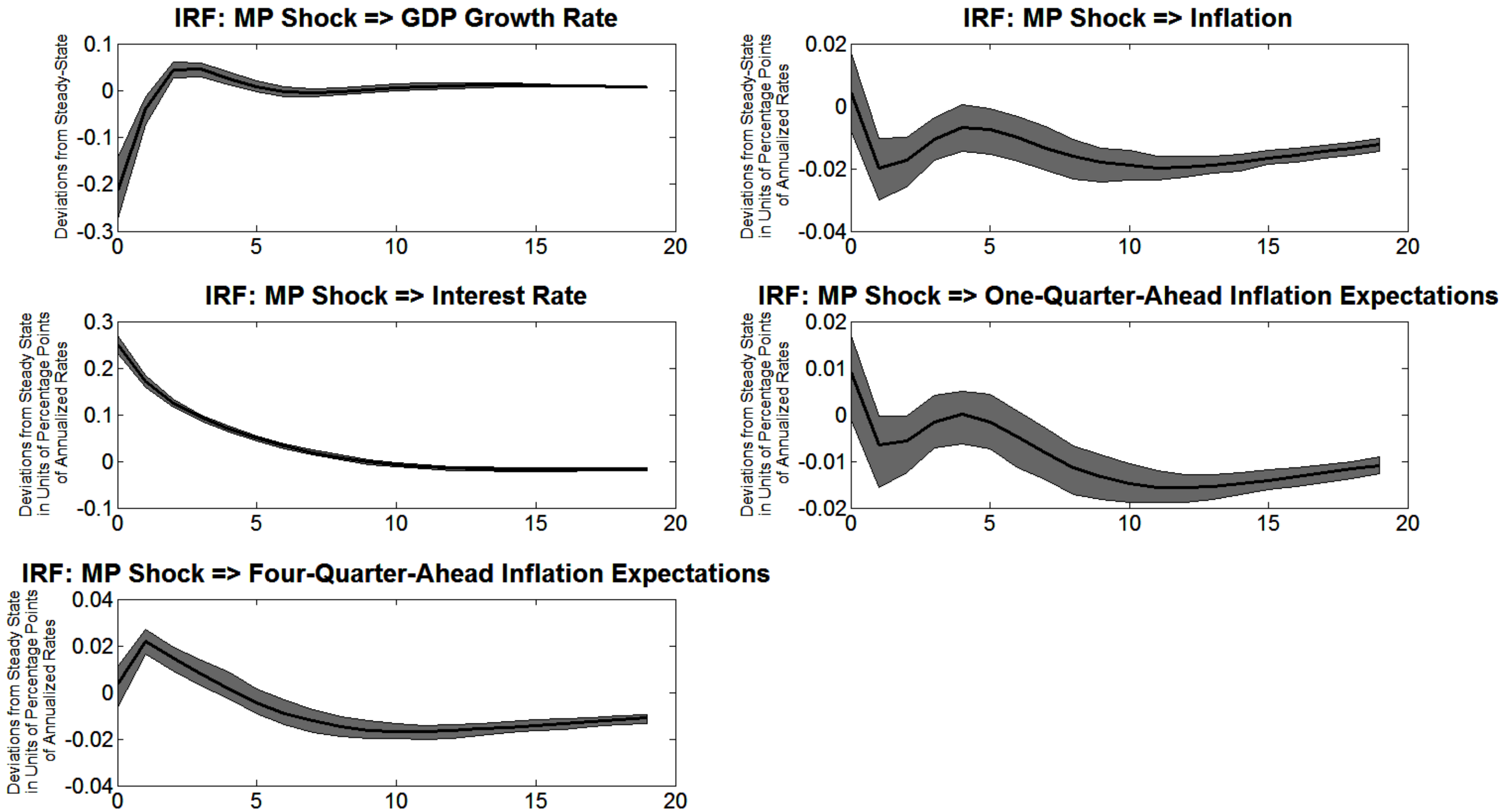

Figure 3: Impulse Response Functions of the Observables to a Monetary Policy Shock. Sample including the 1970s

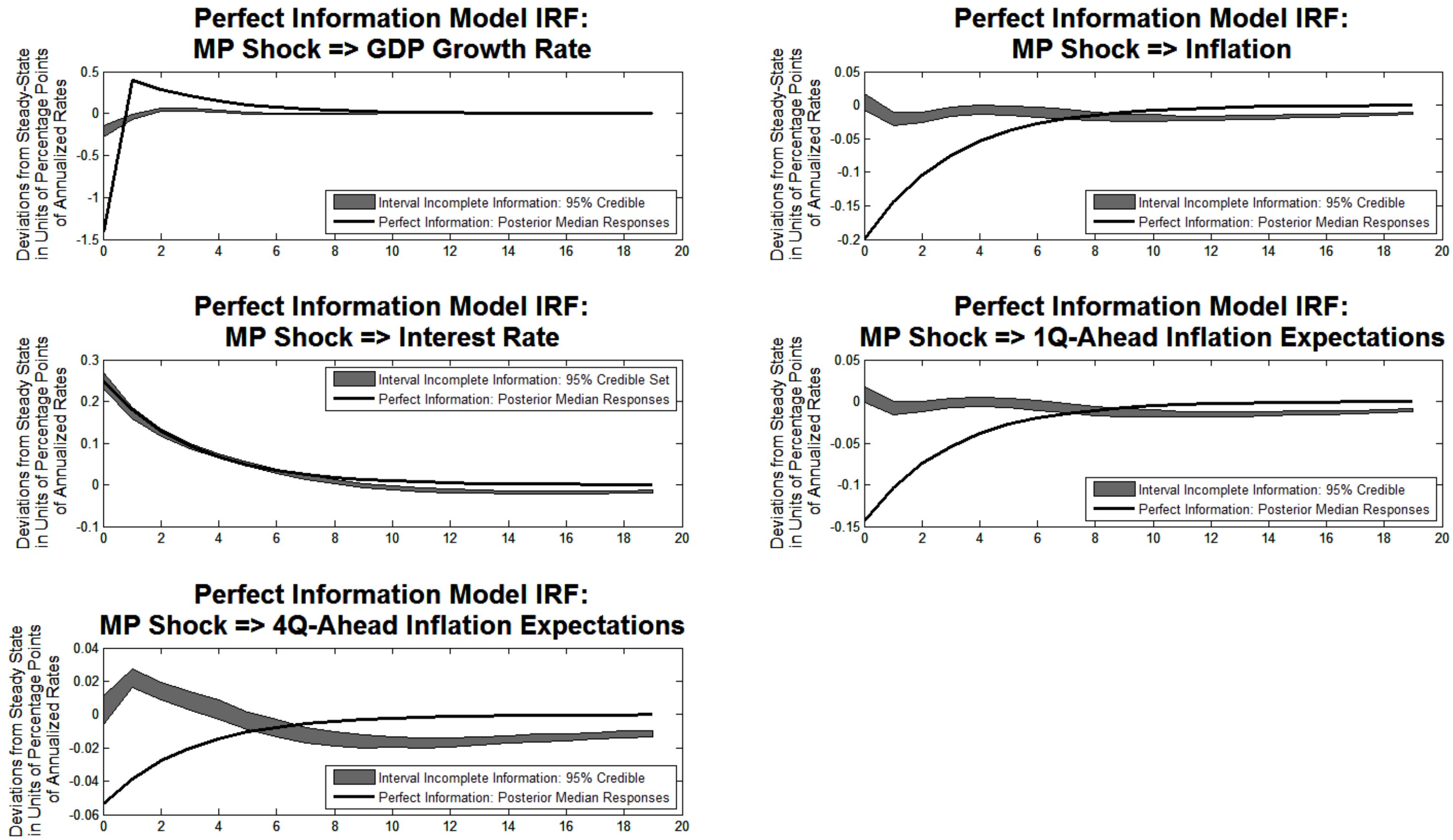

Figure 4: Impulse Response Functions of the Observables to a Monetary Policy Shock: Perfect Information Model (PIM) vs. Incomplete Information Model (IIM). The PIM is calibrated at the posterior medians obtained by estimating the IIM on the sample including the 1970s.

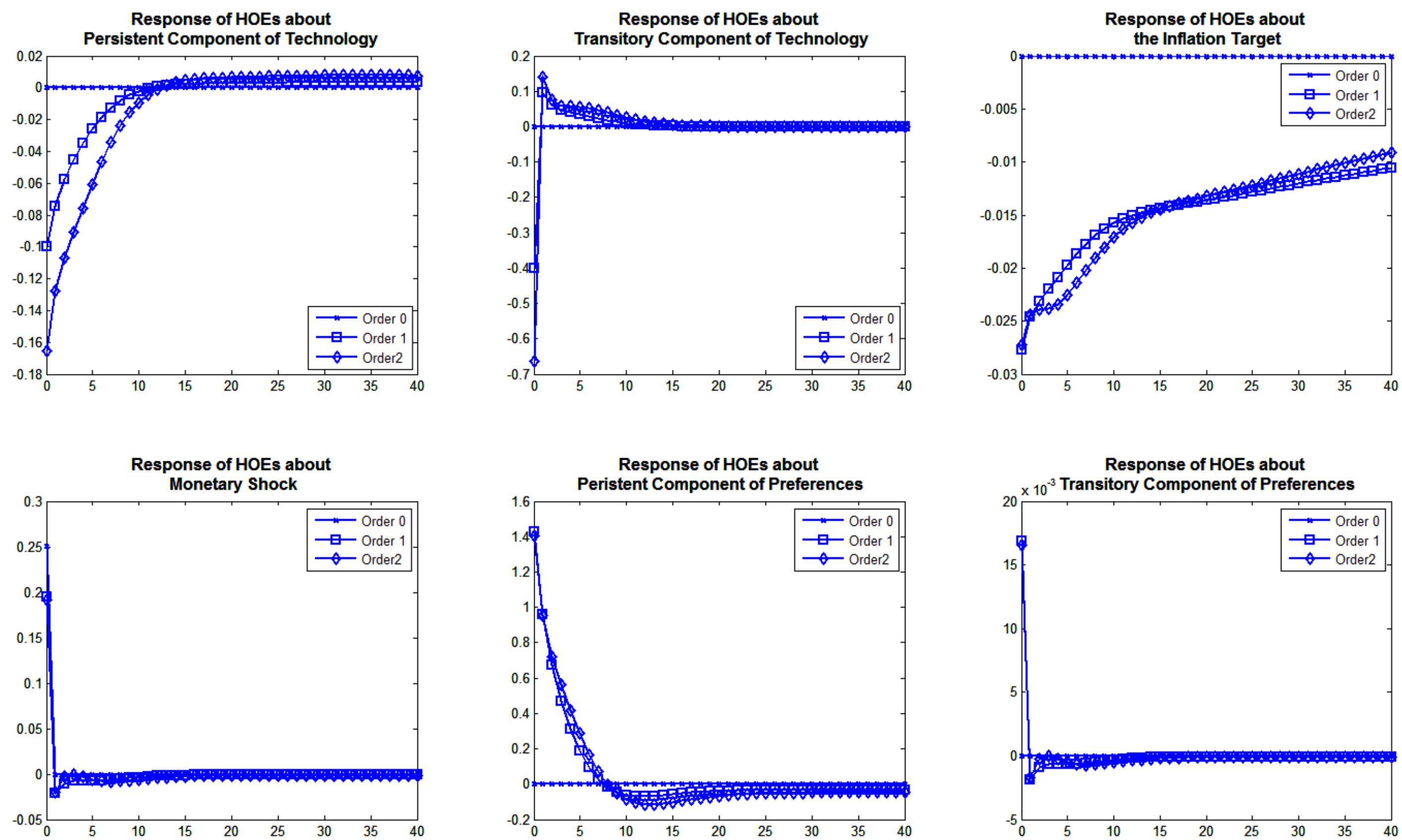

Figure 5: Impulse Response Functions of the High-Order-Expectations about Exogenous Variables to a Monetary Policy Shock. Sample including the 1970s

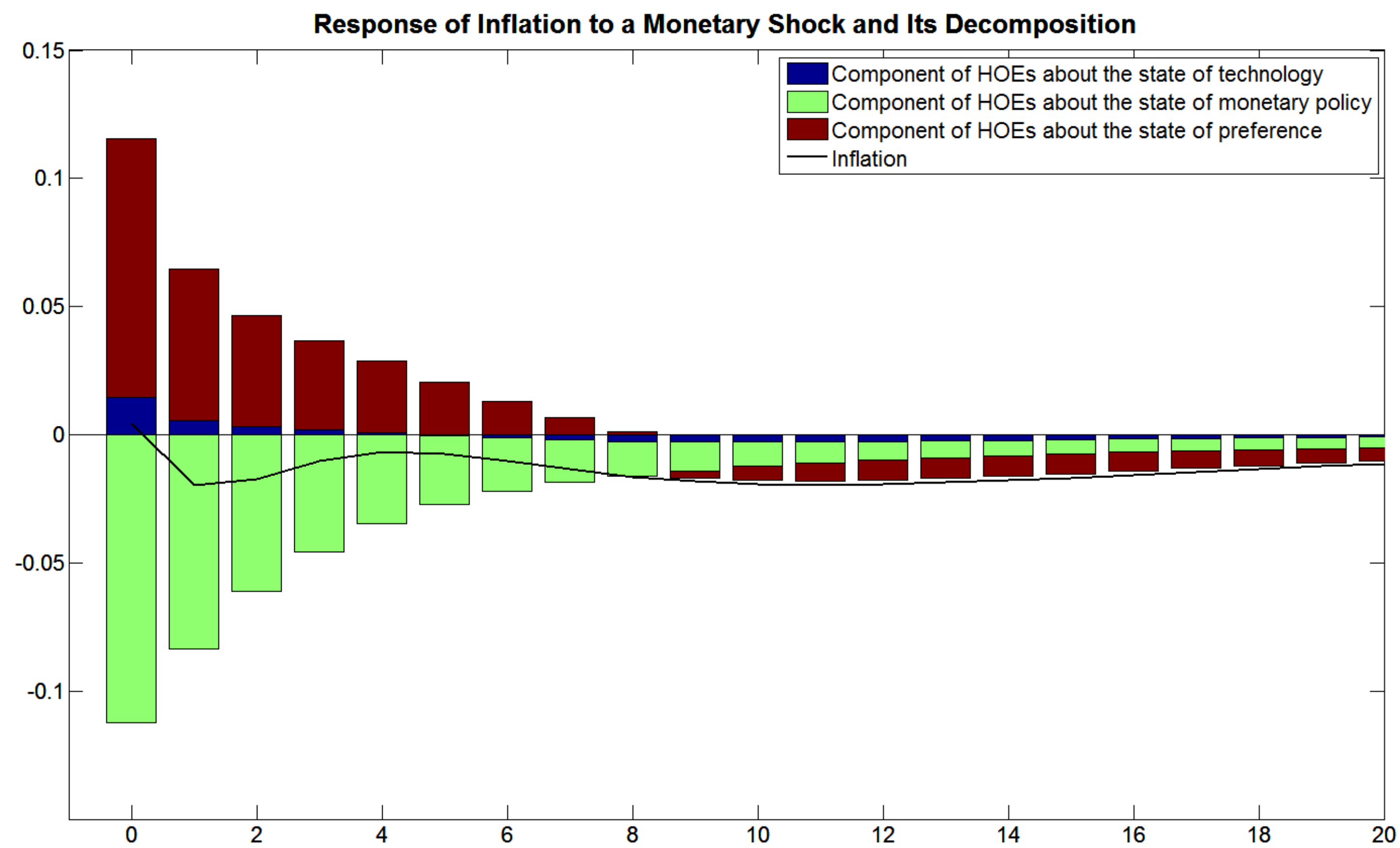

Figure 6: HOE Decomposition of the IRFs of Inflation to a Monetary Policy Shock. Sample including the 1970s

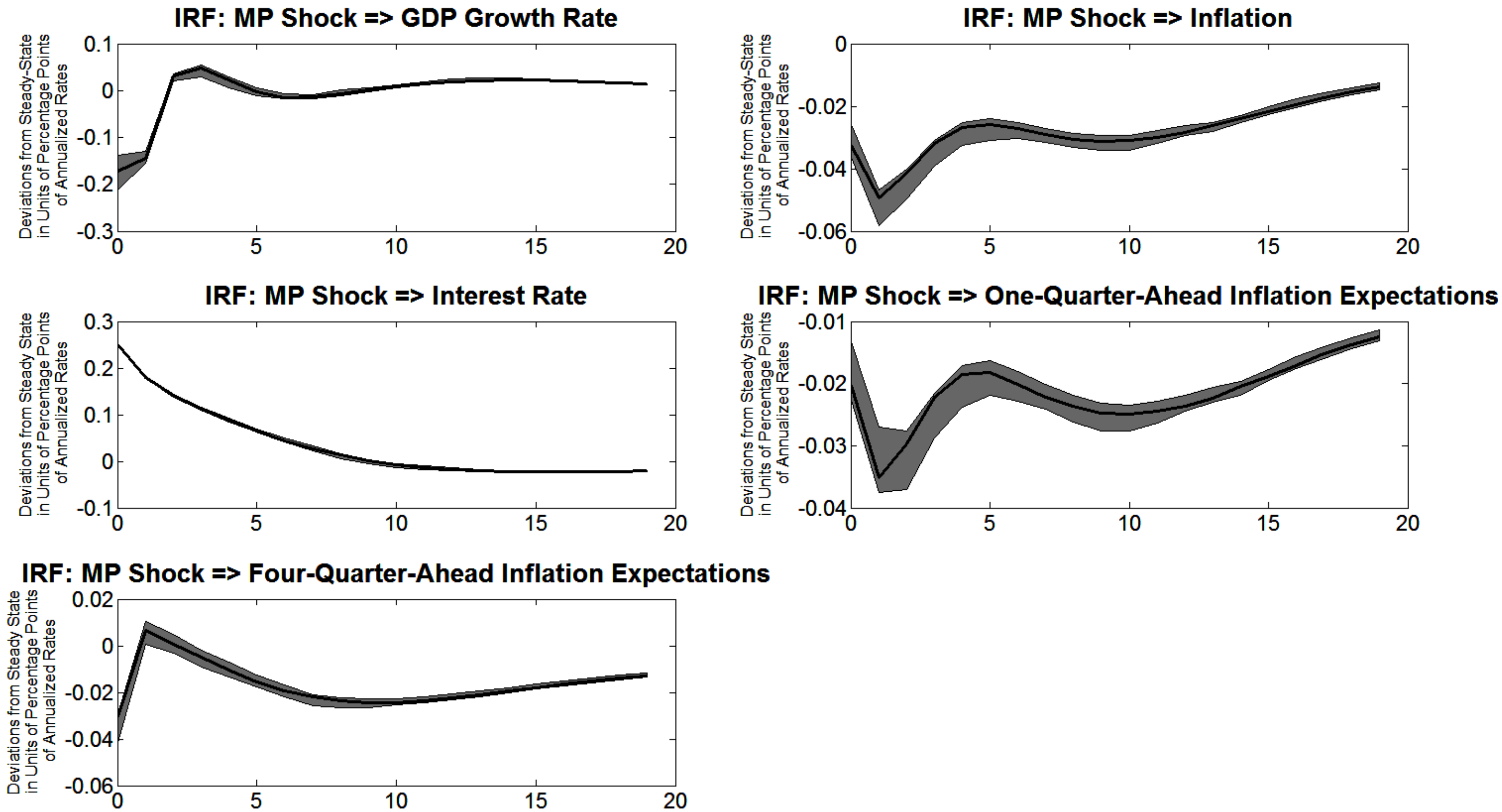

Figure 7: Impulse Response Functions of the Observables to a Monetary Policy Shock in the Sub-samples not including the 1970s

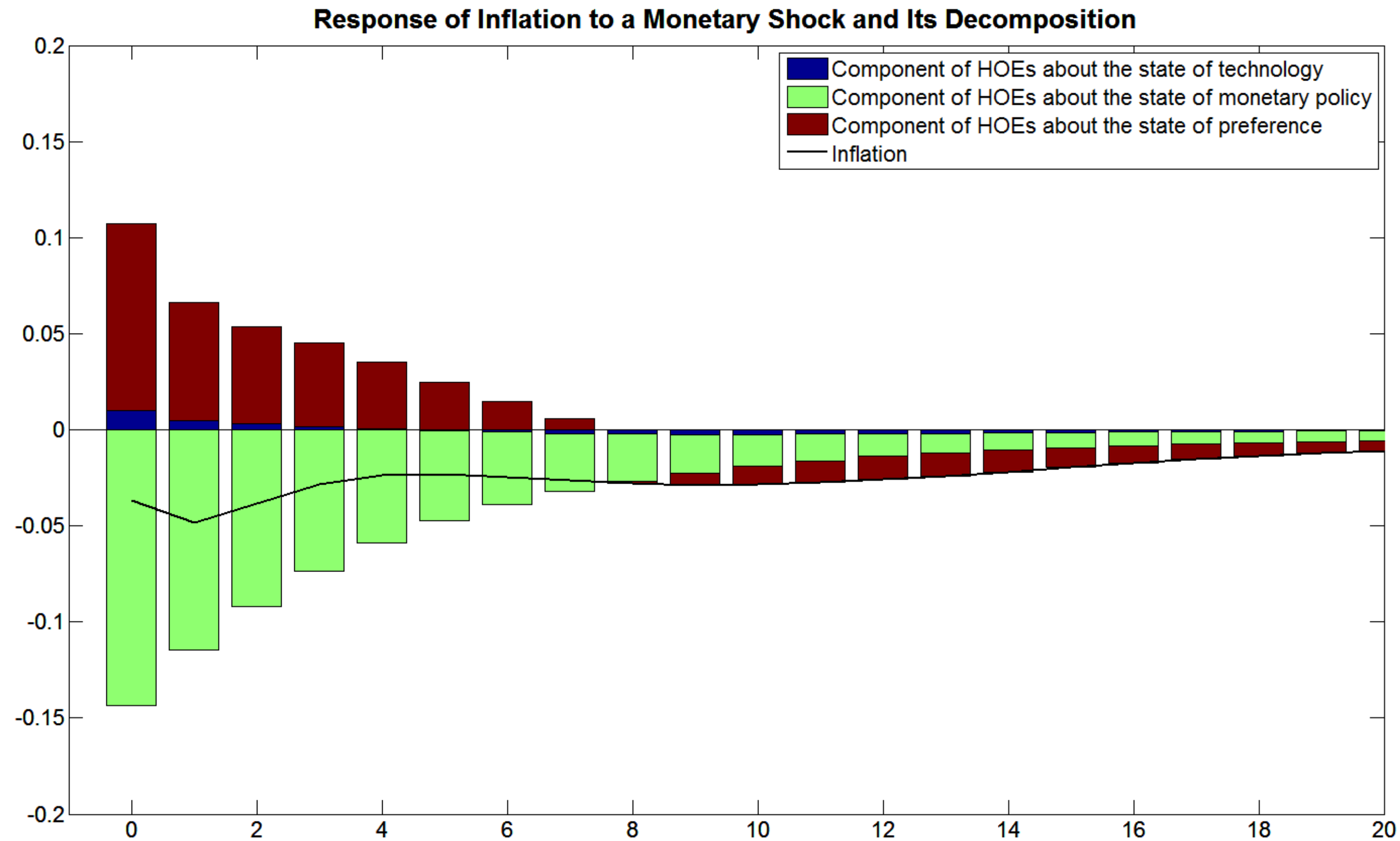

Figure 8: HOE Decomposition of the IRFs of Inflation to a Monetary Policy Shock. Sample not including the 1970s

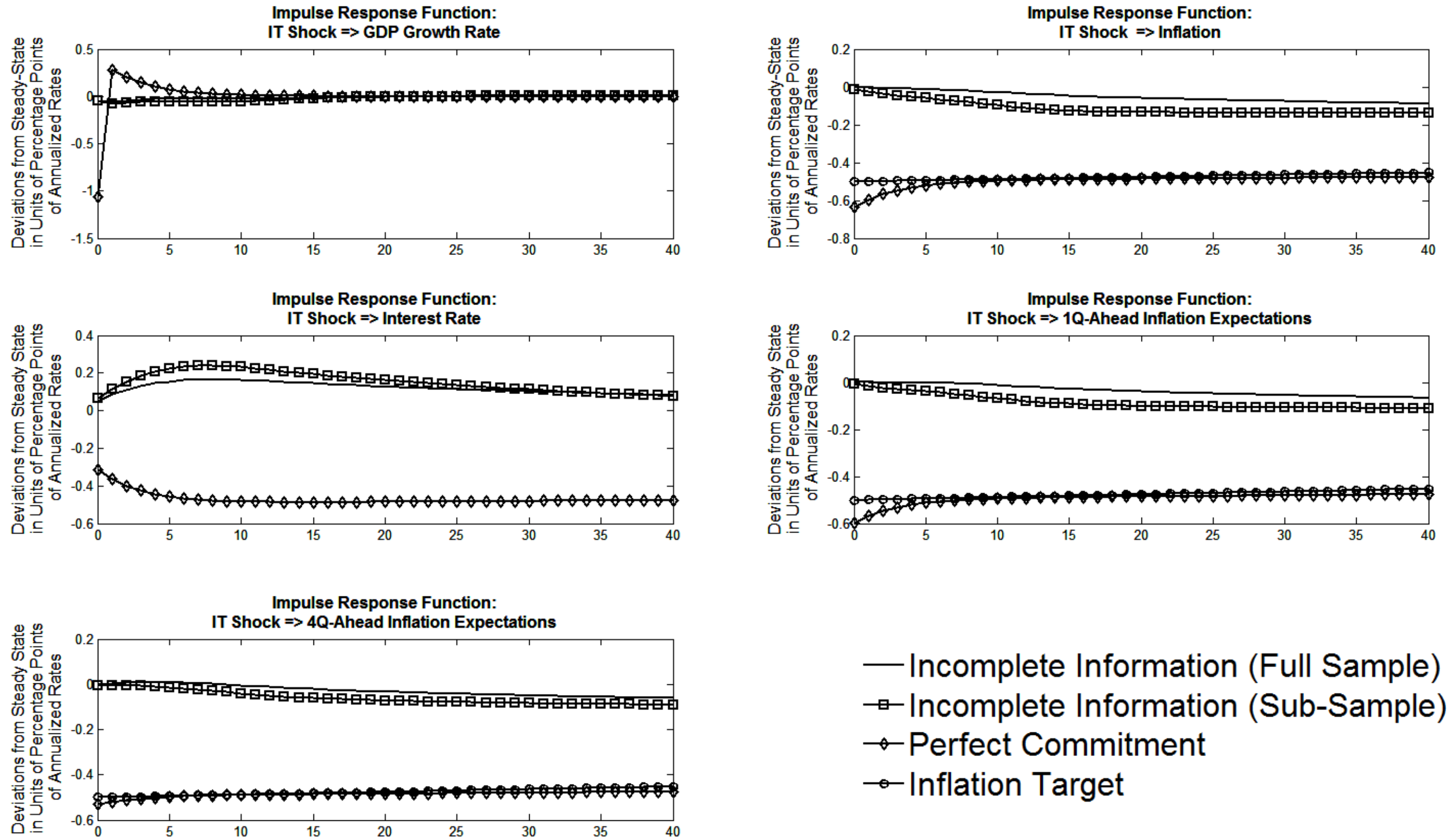

Figure 9: Impulse Response Functions of the Observables to an Inflation-Target Shock: Perfect Commitment, Lack of Commitment in a sample including the 1970s, and Lack of Commitment in a sample not including the 1970s

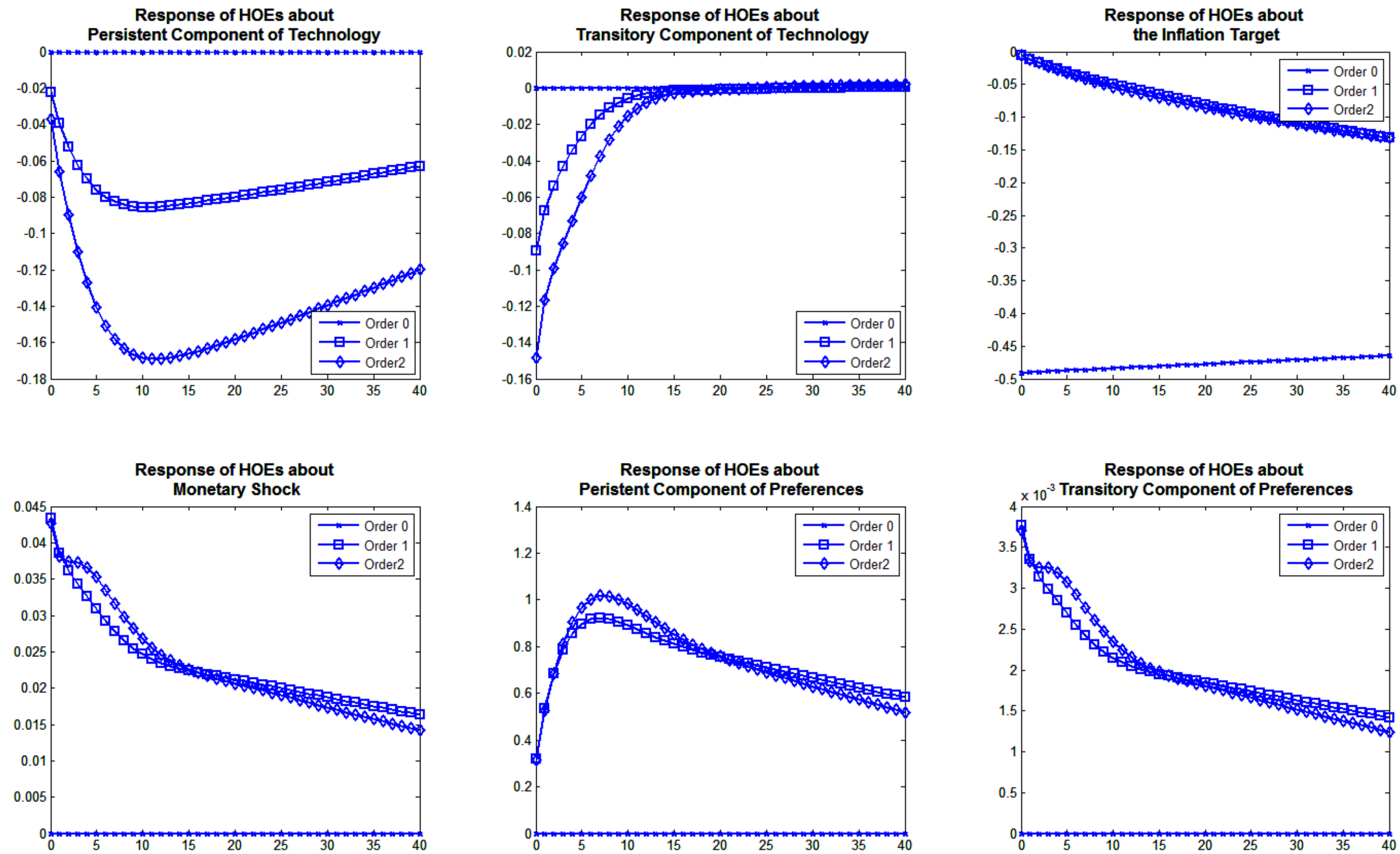

Figure 10: Impulse Response Functions of the High-Order-Expectations about Exogenous Variables to an Inflation-Target Shock (Sample including the 1970s)

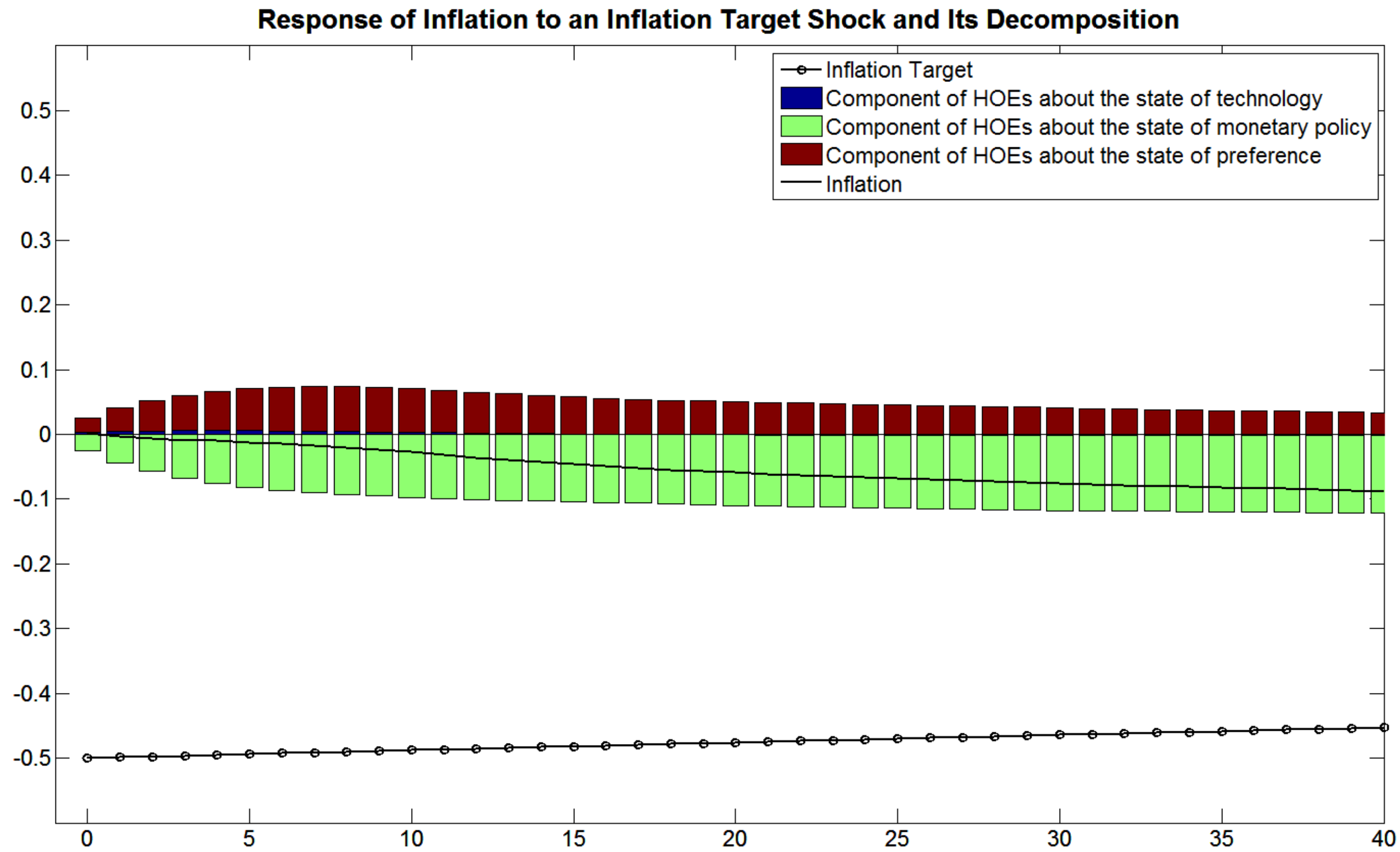

Figure 11: HOE Decomposition of the IRFs of Inflation to an Inflation-Target Shock. Sample including the 1970s

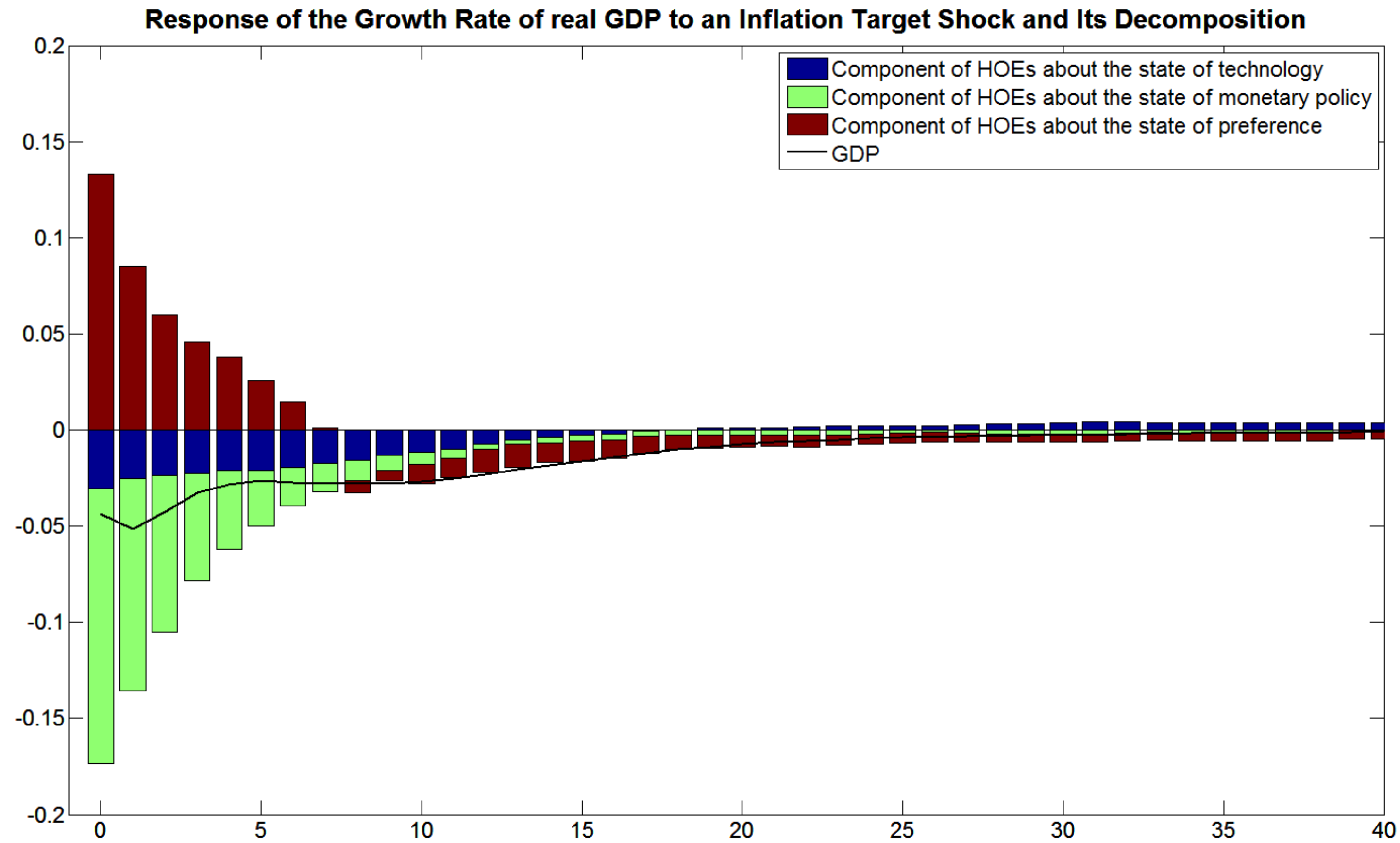

Figure 12: HOE Decomposition of the IRFs of the Growth Rate of GDP to an Inflation-Target Shock. Sample including the 1970s

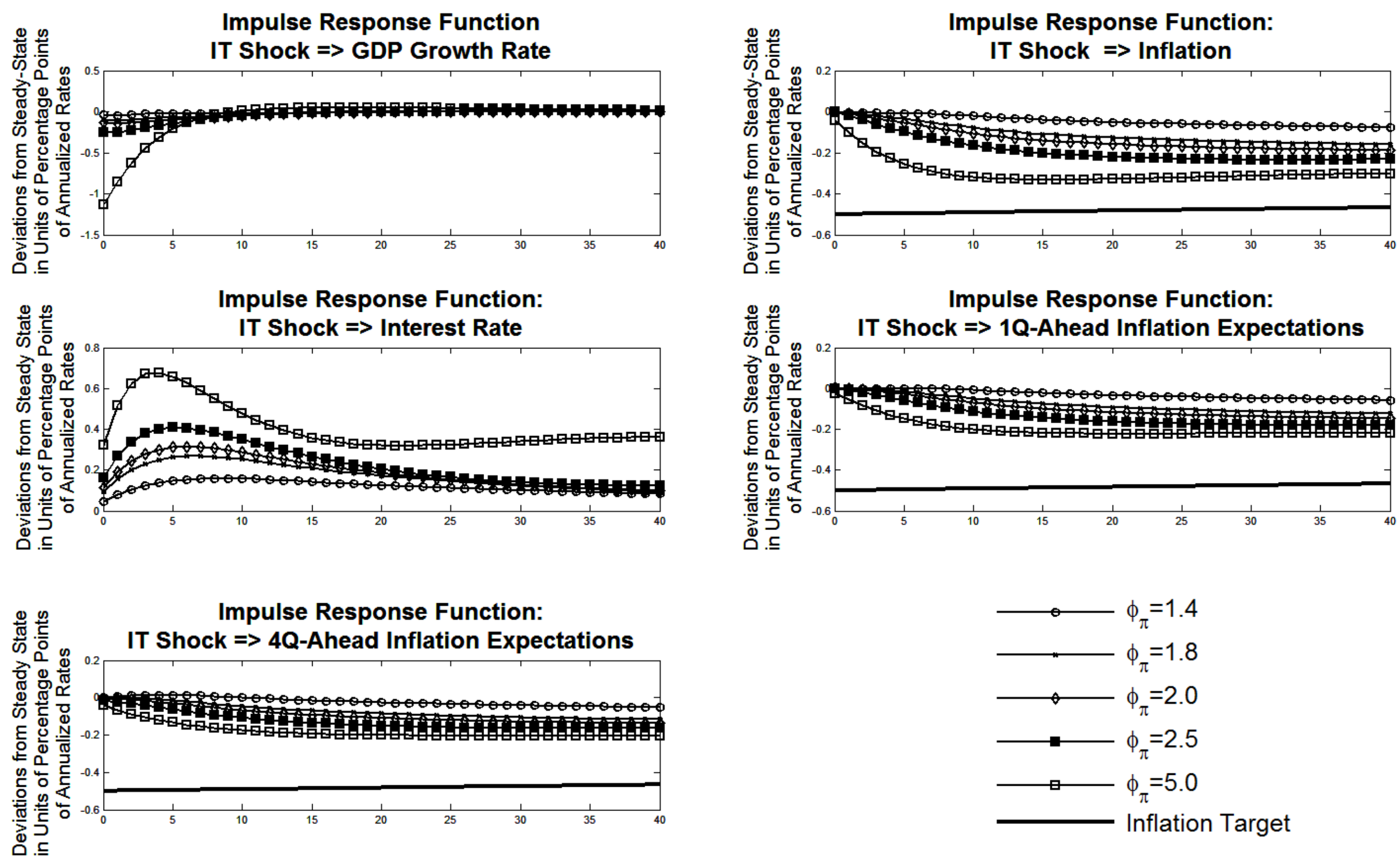

Figure 13: Counterfactual: Impulse Response Function of the GDP, Inflation, Nominal Interest Rate, One-Quarter-Ahead Inflation Expectations, Four-Quarter-Ahead Inflation Expectations under Several Responses to Inflation

Appendix

In Section A, I provide derivation of the imperfect-common-knowledge Phillips curve (11). In Section B, I show how to characterize the coefficients matrices $\mathbf{a}_0$, $\mathbf{b}_0$, $\mathbf{c}_0$, $\mathbf{a}_1$, $\mathbf{b}_1$, $\mathbf{c}_1$ in equations (16)-(18),which are the laws of motion for the three endogenous state variables (i.e., inflation $\hat{\pi}_t$, real output $\hat{y}_t$ and the interest rate $\hat{R}_t$). In Section C, I characterize the coefficient matrices, $\mathbf{M}$ and $\mathbf{N}$, of the equation (15), which is the transition equation for the average higher-order expectations about the exogenous state variables.

A The imperfect common knowledge Phillips curve

The log-linear approximation of the labor supply can be shown to be given by $\widehat{c}_t = \widehat{w}_t$. Recalling that the resource constraint implies that $\hat{y}_t = \hat{c}_t$, then the labor supply can be written as follows:

$$\hat{y}_t = \hat{w}_t \tag{27}$$

Log-linearizing the equation for the real marginal costs (4) yields

$$\widehat{mc}_{j,t} = \hat{w}_{j,t} - z_t - \sigma_a \eta_{a,t} - \eta^a_{j,t}$$

Recall that $(\ln A_{j,t} - \ln A_0 \cdot \gamma t) \in \mathcal{I}_{j,t}$ and write:

$$\mathbb{E}_{j,t}\widehat{mc}_{j,t} = \mathbb{E}_{j,t}\hat{w}_{j,t} \underbrace{- z_t - \sigma_a \eta_{a,t} - \eta^a_{j,t}}_{\ln A_{j,t} - \ln A_0 \cdot \gamma t}$$

where $\mathbb{E}_{j,t}$ are expectations conditioned on firm j's information set at time t, $\mathcal{I}_{j,t}$, defined in (8). Using the equation (27) for replacing $\hat{w}_t$ yields:

$$\mathbb{E}_{j,t}\widehat{mc}_{j,t} = \mathbb{E}_{j,t}\hat{y}_t - z_t - \sigma_a \eta_{a,t} - \eta^a_{j,t}$$

By integrating across firms, we obtain the average expectations on marginal costs:

$$\widehat{mc}^{(0)}_{t|t} = \hat{y}^{(1)}_{t|t} - z_t - \sigma_a \eta_{a,t}$$

The linearized price index can be written as:

$$0 = -\theta \hat{\pi}_t + (1-\theta) \int \hat{p}^*_{j,t} dj$$

By rearranging:

$$\int \hat{p}^*_{j,t} dj = \frac{\theta}{1-\theta} \hat{\pi}_t$$

Recall that we defined $\hat{p}^*_{j,t} = \ln P^*_{j,t} - \ln P_t$ and $\hat{\pi}_t = \ln P_t - \ln P_{t-1} - \ln \pi_*$,

$$\int \ln P^*_{j,t} dj - \ln P_t = \frac{\theta}{1-\theta} (\ln P_t - \ln P_{t-1} - \ln \pi_*)$$

and then

$$\int \ln P^*_{j,t} dj = \frac{1}{1-\theta} \ln P_t - \frac{\theta}{1-\theta} (\ln P_{t-1} + \ln \pi_*)$$

By rearranging:

$$\ln P_t = \theta \left(\ln P_{t-1} + \ln \pi_* \right) + (1-\theta) \int \left(\ln P^*_{j,t} \right) dj \tag{28}$$

The price-setting equation is:

$$\mathbb{E} \left[\sum_{s=0}^{\infty} (\beta\theta)^s \, \Xi_{j,t+s} \left[(1-\nu)\, \pi_*^s + \nu \frac{MC_{j,t+s}}{P^*_{j,t}} \right] Y_{j,t+s} | \mathcal{I}_{j,t} \right] = 0$$

Define

$$\begin{aligned} y_t &= \frac{Y_t}{A_0^t},\ c_t = \frac{C_t}{A_0^t};\ p^*_{j,t} = \frac{P^*_{j,t}}{P_t},\ y_{j,t} = \frac{Y_{j,t}}{A_0^t} \\ w_t &= \frac{W_t}{A_0^t P_t},\ a_t = \frac{A_t}{A_0^t}, R_t = \frac{R_t}{R_*},\ mc_{j,t} = \frac{MC_{j,t}}{P_t} \\ \xi_{j,t} &= A_0^t \Xi_{j,t} \end{aligned}$$

Hence, write:

$$\mathbb{E} \left\{ \xi_{j,t} \left[1 - \nu + \nu \frac{mc_{j,t}}{p^*_{j,t}} \right] y_{j,t} + \sum_{s=1}^{\infty} (\beta\theta)^s \, \xi_{j,t+s} \left[(1-\nu)\, \pi_*^s + \nu \frac{mc_{j,t+s}}{p^*_{j,t}} \left(\Pi_{\tau=1}^s \pi_{t+\tau} \right) \right] y_{j,t+s} | \mathcal{I}_{j,t} \right\} = 0 \tag{29}$$

First realize that the square brackets are equal to zero at the steady state and hence we do not care about the terms outside them. We can write

$$\mathbb{E} \left[\left[1 - \nu + \nu mc_{j,*} e^{\widehat{mc}_{j,t} - \widehat{p}^*_{j,t}} \right] + \sum_{s=1}^{\infty} (\beta\theta)^s \left[(1-\nu) + \nu mc_{j,*} e^{\widehat{mc}_{j,t+s} - \widehat{p}^*_{j,t} + \sum_{\tau=1}^s \hat{\pi}_{t+\tau}} \right] | \mathcal{I}_{j,t} \right] = 0$$

Taking the derivatives yield:

$$\mathbb{E} \left[\widehat{mc}_{j,t} - \widehat{p}^*_{j,t} + \sum_{s=1}^{\infty} (\beta\theta)^s \left[\left(\widehat{mc}_{j,t+s} - \widehat{p}^*_{j,t} + \sum_{\tau=1}^{s} \hat{\pi}_{t+\tau} \right) \right] | \mathcal{I}_{j,t} \right] = 0$$

We can take the term $\widehat{p}_{j,t}$ out of the sum operator in the second term and gather the common term to obtain:

$$\mathbb{E} \left[\widehat{mc}_{j,t} - \frac{1}{1-\beta\theta} \widehat{p}^*_{j,t} + \sum_{s=1}^{\infty} (\beta\theta)^s \left(\widehat{mc}_{j,t+s} + \sum_{\tau=1}^{s} \hat{\pi}_{t+\tau} \right) | \mathcal{I}_{j,t} \right] = 0$$

Recall that $\widehat{p}^*_{j,t} \equiv \ln P^*_{j,t} - \ln P_t$ and cannot be taken out of the expectation operator. We write:

$$\ln P^*_{j,t} = (1-\beta\theta)\, \mathbb{E} \left[\widehat{mc}_{j,t} + \frac{1}{1-\beta\theta} \ln P_t + \sum_{s=1}^{\infty} (\beta\theta)^s \left(\widehat{mc}_{j,t+s} + \sum_{\tau=1}^{s} \hat{\pi}_{t+\tau} \right) | \mathcal{I}_{j,t} \right] \tag{30}$$

Rolling this equation one step ahead yields:

$$\ln P^*_{j,t+1} = (1-\beta\theta)\, \mathbb{E} \left[\widehat{mc}_{j,t+1} + \frac{1}{1-\beta\theta} \ln P_{t+1} + \sum_{s=1}^{\infty} (\beta\theta)^s \left(\widehat{mc}_{j,t+s+1} + \sum_{\tau=1}^{s} \hat{\pi}_{t+\tau+1} \right) | \mathcal{I}_{j,t+1} \right]$$

Take firm j's conditional expectation at time t on both sides and apply the law of iterated expectations:

$$\mathbb{E}\left(\ln P^{*}_{j,t+1}|\mathcal{I}_{j,t}\right) = (1-\beta\theta)\,\mathbb{E}\left[\widehat{mc}_{j,t+1} + \frac{1}{1-\beta\theta}\ln P_{t+1} + \sum_{s=1}^{\infty}(\beta\theta)^{s}\left(\widehat{mc}_{j,t+s+1} + \sum_{\tau=1}^{s}\hat{\pi}_{t+\tau+1}\right)|\mathcal{I}_{j,t}\right]$$

We can take $\widehat{mc}_{j,t+1}$ inside the sum operator and write:

$$\mathbb{E}\left(\ln P^{*}_{j,t+1}|\mathcal{I}_{j,t}\right) = (1-\beta\theta)\,\mathbb{E}\left[\frac{1}{1-\beta\theta}\ln P_{t+1} + \frac{1}{\beta\theta}\sum_{s=1}^{\infty}(\beta\theta)^{s}\,\widehat{mc}_{j,t+s} + \sum_{s=1}^{\infty}(\beta\theta)^{s}\sum_{\tau=1}^{s}\hat{\pi}_{t+\tau+1}|\mathcal{I}_{j,t}\right]$$

Therefore,

$$\sum_{s=1}^{\infty}(\beta\theta)^{s}\,\mathbb{E}\left[\widehat{mc}_{j,t+s}|\mathcal{I}_{j,t}\right] = \frac{\beta\theta}{1-\beta\theta}\left[\mathbb{E}\left(\ln P^{*}_{j,t+1}|\mathcal{I}_{j,t}\right) - \mathbb{E}\left(\ln P_{t+1}|\mathcal{I}_{j,t}\right)\right] - \beta\theta\sum_{s=1}^{\infty}(\beta\theta)^{s}\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau+1}|\mathcal{I}_{j,t}\right] \tag{31}$$

The equation (30) can be rewritten as:

$$\begin{aligned}\ln P^{*}_{j,t} &= (1-\beta\theta)\left\{\mathbb{E}\left[\widehat{mc}_{j,t}|\mathcal{I}_{j,t}\right] + \frac{1}{1-\beta\theta}\mathbb{E}\left[\ln P_{t}|\mathcal{I}_{j,t}\right] + \sum_{s=1}^{\infty}(\beta\theta)^{s}\,\mathbb{E}\left[\widehat{mc}_{j,t+s}|\mathcal{I}_{j,t}\right]\right\}\\ &\quad + (1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s}\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau}|\mathcal{I}_{j,t}\right]\end{aligned}$$

By substituting the result in equation (31) we obtain:

$$\begin{aligned}\ln P^{*}_{j,t} &= (1-\beta\theta)\left[\mathbb{E}\left[\widehat{mc}_{j,t}|\mathcal{I}_{j,t}\right] + \frac{1}{1-\beta\theta}\mathbb{E}\left[\ln P_{t}|\mathcal{I}_{j,t}\right]\right]\\ &\quad +\beta\theta\left[\mathbb{E}\left(\ln P^{*}_{j,t+1}|\mathcal{I}_{j,t}\right) - \mathbb{E}\left(\ln P_{t+1}|\mathcal{I}_{j,t}\right)\right] - (1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s+1}\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau+1}|\mathcal{I}_{j,t}\right]\\ &\quad + (1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s}\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau}|\mathcal{I}_{j,t}\right]\end{aligned}$$

Consider the last term:

$$\begin{aligned}(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s}\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau}|\mathcal{I}_{j,t}\right] &= (1-\beta\theta)\,\beta\theta\mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right] + (1-\beta\theta)\sum_{s=2}^{\infty}(\beta\theta)^{s}\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau}|\mathcal{I}_{j,t}\right]\\ &= (1-\beta\theta)\,\beta\theta\mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right] +\\ &\quad + (1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s+1}\left(\sum_{\tau=1}^{s}\left[\left(\mathbb{E}\left[\hat{\pi}_{t+\tau+1}|\mathcal{I}_{j,t}\right]\right)\right] + \mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right]\right)\end{aligned}$$

Therefore we can write that

$$\begin{aligned}(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s}\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau}|\mathcal{I}_{j,t}\right] &= (1-\beta\theta)\,\beta\theta\mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right]\\ &+(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s+1}\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau+1}|\mathcal{I}_{j,t}\right]\\ &+(1-\beta\theta)\left(\sum_{s=1}^{\infty}(\beta\theta)^{s+1}\right)\mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right]\end{aligned}$$

Note that

$$\left(\sum_{s=1}^{\infty}(\beta\theta)^{s+1}\right)=\frac{(\beta\theta)^{2}}{1-\beta\theta}$$

Hence,

$$\begin{aligned}(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s}\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau}|\mathcal{I}_{j,t}\right] &= (1-\beta\theta)\,\beta\theta\mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right]\\ &+(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s+1}\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau+1}|\mathcal{I}_{j,t}\right]\\ &+(\beta\theta)^{2}\,\mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right]\end{aligned}$$

and by simplifying:

$$\begin{aligned}(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s}\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau}|\mathcal{I}_{j,t}\right] &= \beta\theta\mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right]\\ &+(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s+1}\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau+1}|\mathcal{I}_{j,t}\right]\end{aligned}$$

We substitute this result into the original equation to get:

$$\begin{aligned}\ln P_{j,t}^{*} &= (1-\beta\theta)\left[\mathbb{E}\left[\widehat{mc}_{j,t}|\mathcal{I}_{j,t}\right]+\frac{1}{1-\beta\theta}\mathbb{E}\left[\ln P_{t}|\mathcal{I}_{j,t}\right]\right]\\ &+\beta\theta\left[\mathbb{E}\left(\ln P_{j,t+1}^{*}|\mathcal{I}_{j,t}\right)-\mathbb{E}\left(\ln P_{t+1}|\mathcal{I}_{j,t}\right)\right]-(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s+1}\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau+1}|\mathcal{I}_{j,t}\right]\\ &+\beta\theta\mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right]+(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s+1}\sum_{\tau=1}^{s}\mathbb{E}\left[\hat{\pi}_{t+\tau+1}|\mathcal{I}_{j,t}\right] \qquad (32)\end{aligned}$$

After simplifying we get:

$$\begin{aligned}\ln P_{j,t}^{*} &= (1-\beta\theta)\left[\mathbb{E}\left[\widehat{mc}_{j,t}|\mathcal{I}_{j,t}\right]+\frac{1}{1-\beta\theta}\mathbb{E}\left[\ln P_{t}|\mathcal{I}_{j,t}\right]\right]\\ &+\beta\theta\left[\mathbb{E}\left(\ln P_{j,t+1}^{*}|\mathcal{I}_{j,t}\right)-\mathbb{E}\left(\ln P_{t+1}|\mathcal{I}_{j,t}\right)\right]+\beta\theta\mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right] \qquad (33)\end{aligned}$$

We can rearrange:

$$\begin{aligned}\ln P^*_{j,t} &= (1-\beta\theta)\,\mathbb{E}\left[\widehat{mc}_{j,t}|\mathcal{I}_{j,t}\right]+\mathbb{E}\left[\ln P_t|\mathcal{I}_{j,t}\right] \\ &\quad+\beta\theta\left[\mathbb{E}\left(\ln P^*_{j,t+1}|\mathcal{I}_{j,t}\right)+\mathbb{E}\left[\hat{\pi}_{t+1}|\mathcal{I}_{j,t}\right]-\mathbb{E}\left(\ln P_{t+1}|\mathcal{I}_{j,t}\right)\right] \end{aligned} \tag{34}$$

Note that by definition $\hat{\pi}_{t+1}\equiv \ln P_{t+1}-\ln P_t-\ln\pi_*$. Hence we can show that

$$\begin{aligned}\ln P^*_{j,t} &= (1-\beta\theta)\cdot\mathbb{E}\left[\widehat{mc}_{j,t}|\mathcal{I}_{j,t}\right]+(1-\beta\theta)\,\mathbb{E}\left[\ln P_t|\mathcal{I}_{j,t}\right] \\ &\quad+\beta\theta\cdot\mathbb{E}\left(\ln P^*_{j,t+1}|\mathcal{I}_{j,t}\right)-\beta\theta\ln\pi_* \end{aligned} \tag{35}$$

We denote the firm j's average k-th order expectation about an arbitrary variable $\hat{x}_t$ as

$$\mathbb{E}^{(k)}\left(\hat{x}_t|\mathcal{I}_{j,t}\right)\equiv\int\mathbb{E}\left(\int\mathbb{E}\left(\ldots\left(\int\mathbb{E}\left(\hat{x}_t|\mathcal{I}_{j,t}\right)dj\right)\ldots|\mathcal{I}_{j,t}\right)dj|\mathcal{I}_{j,t}\right)dj$$

where expectations and integration across firms are taken k times.

Let us denote the **average reset price** as $\ln P^*_t=\int\ln P^*_{j,t}dj$. We can integrate equation (35) across firms to obtain an equation for the average reset price:

$$\begin{aligned}\ln P^*_t &= (1-\beta\theta)\cdot\widehat{mc}^{(0)}_{t|t}+(1-\beta\theta)\ln P^{(1)}_{t|t} \\ &\quad+\beta\theta\ln P^{*(1)}_{t+1|t}-\beta\theta\ln\pi_* \end{aligned} \tag{36}$$

where we use the claim of the proposition above. Keep in mind that the price index equation can be manipulated to get equation (28)

$$\ln P_t=\theta\left(\ln P_{t-1}+\ln\pi_*\right)+(1-\theta)\ln P^*_t \tag{37}$$

Let us plug the equation (36) into the equation (37):

$$\begin{aligned}\ln P_t &= \theta\ln P_{t-1}+\left(\theta-(1-\theta)\,\beta\theta\right)\ln\pi_* \\ &\quad+(1-\theta)\left[(1-\beta\theta)\cdot\widehat{mc}^{(0)}_{t|t}+(1-\beta\theta)\ln P^{(1)}_{t|t}+\beta\theta\ln P^{*(1)}_{t+1|t}\right] \end{aligned} \tag{38}$$

Use the fact that $\ln P_t = \widehat{\pi}_t + \ln P_{t-1} + \ln \pi_*$ and from the price index (28):[22]

$$\ln P^*_{t+1} = \frac{\hat{\pi}_{t+1}}{1-\theta} + \ln P_t + \ln \pi_*$$

Furthermore, the following fact is easy to establish:

$$\ln P_{t+1} = \widehat{\pi}_{t+1} + \ln P_t + \ln \pi_*$$

Applying these three results to equation (38) yields:

$$\begin{aligned} \widehat{\pi}_t + \ln P_{t-1} + \ln \pi_* \quad = \quad & \theta \ln P_{t-1} + (\theta - (1-\theta)\,\beta\theta) \ln \pi_* \qquad (39) \\ & + (1-\theta) \left[(1-\beta\theta) \cdot \widehat{mc}^{(0)}_{t|t} + (1-\beta\theta) \ln P^{(1)}_{t|t} + \beta\theta \left(\frac{\hat{\pi}^{(1)}_{t+1|t}}{1-\theta} + \ln P^{(1)}_{t|t} + \ln \pi_* \right) \right] \end{aligned}$$

and after simplifying:

$$\widehat{\pi}_t = (1-\theta)\,(1-\beta\theta) \cdot \widehat{mc}^{(0)}_{t|t} + (1-\theta)\,\widehat{\pi}^{(1)}_{t|t} + \beta\theta \left(\hat{\pi}^{(1)}_{t+1|t} \right) \qquad (40)$$

By repeatedly taking firm j's expectation and average the resulting equation across firms:

$$\widehat{\pi}^{(k)}_{t|t} = (1-\theta)\,(1-\beta\theta) \cdot \widehat{mc}^{(k)}_{t|t} + (1-\theta)\,\widehat{\pi}^{(k+1)}_{t|t} + \beta\theta \left(\hat{\pi}^{(k+1)}_{t+1|t} \right)$$

Repeatedly substituting these equations for $k \geq 1$ back to equation (40) yields: the imperfect-common-knowledge Phillips curve:

$$\widehat{\pi}_t = (1-\theta)\,(1-\beta\theta) \sum_{k=0}^{\infty} (1-\theta)^k\, \widehat{mc}^{(k)}_{t|t} + \beta\theta \sum_{k=0}^{\infty} (1-\theta)^k\, \hat{\pi}^{(k+1)}_{t+1|t} \qquad (41)$$

B The Laws of Motion for the Endogenous State Variables

In this section I, first, introduce some useful results and, second, characterize the law of motion for the endogenous state variables of the model (16)-(18), which are inflation $\hat{\pi}_t$, real output $\hat{y}_t$, and the (nominal) interest rate $\hat{R}_t$. It will be shown that this law of motion depends on model parameters

[22] This last result comes from observing that

$$\ln P_t = \theta\,(\ln P_{t-1} + \ln \pi_*) + (1-\theta) \ln P^*_t$$

By using the fact that $\ln P_t = \widehat{\pi}_t + \ln P_{t-1} + \ln \pi_*$:

$$\widehat{\pi}_t + \ln P_{t-1} + \ln \pi_* = \theta\,(\ln P_{t-1} + \ln \pi_*) + (1-\theta) \ln P^*_t$$

Rolling one period forward:

$$\widehat{\pi}_{t+1} = (\theta - 1)\,(\ln P_t + \ln \pi_*) + (1-\theta) \ln P^*_{t+1}$$

and finally by rearranging we get the result in the text.

and the coefficient matrices, $\mathbf{M}$ and $\mathbf{N}$, of the transition equation for the average higher-order expectations about the exogenous variables.

B.1 Preliminaries

Recall that the *assumption of common knowledge in rationality* ensures that agents use the actual law of motion of higher-order expectations to forecast the dynamics of the higher-order expectations. The following claims turn out to be useful:

Proposition 1 *If one neglects the effect of average beliefs of order larger than k, then the following is approximately true:*

$$\varphi_{t|t}^{(s:k+s)} = \mathbf{T}^{(s)} \varphi_{t|t}^{(0:k)}$$

where

$$\mathbf{T}^{(s)} \equiv \begin{bmatrix} \mathbf{0}_{6(k-s+1)\times 6s} & \mathbf{I}_{6(k-s+1)} \\ \mathbf{0}_{6s\times 6s} & \mathbf{0}_{6s\times(k+1-s)6} \end{bmatrix}$$

Proof. It is straightforward but help to fix some notation. Since we neglect the average beliefs of order larger than k

$$\varphi_{t|t}^{(s:k+s)} \equiv \begin{bmatrix} \varphi_{t|t}^{(s:k)} \\ \varphi_{t|t}^{(s:k+s)} \end{bmatrix}_{6(k+1)\times 1} = \begin{bmatrix} \varphi_{t|t}^{(s:k)} \\ \mathbf{0}_{6s\times 1} \end{bmatrix}_{6(k+1)\times 1}$$

Note that

$$\varphi_{t|t}^{(s:k+s)} = \begin{bmatrix} \mathbf{0}_{6(k-s+1)\times 6s} & \mathbf{I}_{6(k-s+1)} \\ \mathbf{0}_{6s\times 6s} & \mathbf{0}_{6s\times(k+1-s)6} \end{bmatrix} \underbrace{\begin{bmatrix} \varphi_{t|t}^{(0:s-1)} \\ \varphi_{t|t}^{(s:k)} \end{bmatrix}}_{\varphi_{t|t}^{(0:k)}}$$

■

Proposition 2 $\mathbf{s}_{t|t}^{(s)} = \mathbf{v}_0 \mathbf{T}^{(s)} \varphi_{t|t}^{(0:k+s)} + \mathbf{v}_1 \mathbf{s}_{t-1}$, *for any* $0 \leq s \leq k$.

Proof. We conjectured that $\mathbf{s}_t = \mathbf{v}_0 \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_{t-1}$. Then common knowledge in rationality implies:

$$\mathbf{s}_{t|t}^{(s)} = \mathbf{v}_0 \varphi_{t|t}^{(s:k+s)} + \mathbf{v}_1 \mathbf{s}_{t-1}$$

■

Since we truncate beliefs after the k-th order we have that

$$\mathbf{s}_{t|t}^{(s)} = \mathbf{v}_0 \mathbf{T}^{(s)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_{t-1}, \text{for any } 0 \leq s \leq k$$

Proposition 3 *The following holds true for any* $h \in \{0 \cup \mathbb{N}\}$

$$\mathbf{s}_{t+h|t}^{(1)} = \sum_{l=0}^{h} \mathbf{v}_1^{h-l} \mathbf{v}_0 \mathbf{M}^l \mathbf{T}^{(1)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1^{h+1} \mathbf{s}_{t-1}$$

Proof. Consider

$$\mathbf{s}_t = \mathbf{v}_0 \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_{t-1}$$

Then it is easy to see that by taking agents' expectations and then averaging across them we obtain by the *assumption of common knowledge in rationality*:

$$\mathbf{s}_{t|t}^{(1)} = \mathbf{v}_0 \varphi_{t|t}^{(1:k+1)} + \mathbf{v}_1 \mathbf{s}_{t-1}$$

and by neglecting the contribution of beliefs of order higher than k we can write: $\mathbf{T}^{(1)} \varphi_{t|t}^{(0:k)} = \varphi_{t|t}^{(1:k+1)}$. This leads to write:

$$\mathbf{s}_{t|t}^{(1)} = \mathbf{v}_0 \mathbf{T}^{(1)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_{t-1} \tag{42}$$

Furthermore, consider $\mathbf{s}_{t+1}$:

$$\mathbf{s}_{t+1} = \mathbf{v}_0 \varphi_{t+1|t+1}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_t$$

By taking agents' expectations and then averaging across them we obtain:

$$\mathbf{s}_{t+1|t}^{(1)} = \mathbf{v}_0 \varphi_{t+1|t}^{(1:k+1)} + \mathbf{v}_1 \mathbf{s}_{t|t}^{(1)}$$

Recall result (42) and write:

$$\mathbf{s}_{t+1|t}^{(1)} = \mathbf{v}_0 \varphi_{t+1|t}^{(1:k+1)} + \mathbf{v}_1 \left[\mathbf{v}_0 \mathbf{T}^{(1)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_{t-1} \right]$$

First note that by the assumption of common knowledge in rationality we can write: $\varphi_{t+h|t}^{(1:k+1)} = \mathbf{M}^h \varphi_{t|t}^{(1:k+1)}$. Second, recall that we neglect the contribution of beliefs of order higher than k. These two facts lead us to

$$\mathbf{s}_{t+1|t}^{(1)} = \mathbf{v}_0 \mathbf{M} \mathbf{T}^{(1)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \left[\mathbf{v}_0 \mathbf{T}^{(1)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_{t-1} \right]$$

Consider now $\mathbf{s}_{t+2}$. By taking agents' expectations and then averaging across them we obtain:

$$\mathbf{s}_{t+2|t}^{(1)} = \mathbf{v}_0 \varphi_{t+2|t}^{(1:k+1)} + \mathbf{v}_1 \mathbf{s}_{t+1|t}^{(1)}$$

and substituting $\mathbf{s}_{t+1|t}^{(1)}$ that we have characterized above yields:

$$\mathbf{s}_{t+2|t}^{(1)} = \mathbf{v}_0 \mathbf{M}^2 \mathbf{T}^{(1)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \left\{ \mathbf{v}_0 \mathbf{M} \mathbf{T}^{(1)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \left[\mathbf{v}_0 \mathbf{T}^{(1)} \varphi_{t|t}^{(0:k)} + \mathbf{v}_1 \mathbf{s}_{t-1} \right] \right\}$$

Keeping on deriving $\mathbf{s}_{t+h|t}^{(1)}$ for any other $h \in \{0 \cup \mathbb{N}\}$ as shown above leads at the formula in the claim. ■

B.2 The Laws of Motion of the Endogenous State Variables

The laws of motion of the three endogenous state variables (16)-(18), which are inflation $\hat{\pi}_t$, real output $\hat{y}_t$, and the (nominal) interest rate $\hat{R}_t$, are given by the IS equation (13), the Phillips curve (11), and the Taylor Rule (14). One can use these structural equations to pin down the vectors $\mathbf{v}_0 \equiv \left[\mathbf{a}_0', \mathbf{b}_0', \mathbf{c}_0'\right]'$ and $\mathbf{v}_1 \equiv \left[\mathbf{a}_1', \mathbf{b}_1', \mathbf{c}_1'\right]'$ in equations (16)-(18).

Let start from the IS equation (13)

$$\begin{aligned}
\mathbf{b}_0\varphi_{t|t}^{(0:k)} + \mathbf{b}_1\mathbf{s}_{t-1} &= \left(\mathbf{1}_5^T + \mathbf{1}_6^T\right)\varphi_{t|t}^{(0:k)} - \left(\mathbf{1}_5^T + \mathbf{1}_6^T\right)\mathbf{M}\varphi_{t|t}^{(0:k)} \\
&+ \left(\mathbf{1}_1^T + \mathbf{1}_2^T\right)\left[\underbrace{\mathbf{v}_0\mathbf{M}\mathbf{T}^{(1)}\varphi_{t|t}^{(0:k)} + \mathbf{v}_1\left(\mathbf{v}_0^{(1)}\mathbf{T}^{(1)}\varphi_{t|t}^{(0:k)} + \mathbf{v}_1\mathbf{s}_{t-1}\right)}_{\mathbf{s}_{t+1|t}^{(1)}}\right] \\
&- \left(\mathbf{c}_0\varphi_{t|t}^{(0:k)} + \mathbf{c}_1\mathbf{s}_{t-1}\right)
\end{aligned}$$

where we used Proposition 3, $\mathbf{1}_i^T$ $(i \in \mathbb{N})$ is a comfortable row vector of all zero elements except for the i-th element, which is equal to one. One can show that the following condition has to be satisfied by the vectors of coefficients, $\mathbf{b}_0$ and $\mathbf{b}_1$, to ensure that the IS equation (13) holds in equilibrium.

$$\mathbf{b}_0 = \left(\mathbf{1}_5^T + \mathbf{1}_6^T\right) + \left(\mathbf{1}_1^T + \mathbf{1}_2^T\right)\left(\mathbf{v}_0\mathbf{M}\mathbf{T}^{(1)} + \mathbf{v}_1\mathbf{v}_0\mathbf{T}^{(1)}\right) - \left(\mathbf{1}_5^T + \mathbf{1}_6^T\right)\mathbf{M} - \mathbf{c}_0 \tag{43}$$

$$\mathbf{b}_1 = \left(\mathbf{1}_1^T + \mathbf{1}_2^T\right)\mathbf{v}_1\mathbf{v}_1 - \mathbf{c}_1 \tag{44}$$

The Phillips curve (11) can be rewritten as:

$$\begin{aligned}
&\mathbf{a}_0\varphi_{t|t}^{(0:k)} + \mathbf{a}_1\mathbf{s}_{t-1} = (1-\theta)(1-\beta\theta)\sum_{s=0}^{k-1}(1-\theta)^s\,\mathbf{1}_2^T\left[\mathbf{v}_0\mathbf{T}^{(s+1)}\varphi_{t|t}^{(0:k)} + \mathbf{v}_1\mathbf{s}_{t-1}\right] + \\
&-(1-\theta)(1-\beta\theta)\sum_{s=0}^{k-1}(1-\theta)^s\left[\boldsymbol{\gamma}_a^{(s)\prime}\varphi_{t|t}^{(0:k)}\right] \\
&+\beta\theta\sum_{s=0}^{k-1}(1-\theta)^s\,\mathbf{1}_1^T\left[\mathbf{v}_0\mathbf{M}\mathbf{T}^{(s+1)}\varphi_{t|t}^{(0:k)} + \mathbf{v}_1\left(\mathbf{v}_0\varphi_{t|t}^{(0:k)} + \mathbf{v}_1\mathbf{s}_{t-1}\right)\right]
\end{aligned}$$

where $\boldsymbol{\gamma}_a^{(s)} = \left[\mathbf{0}_{1\times 6s}, (1,1,0,0,0,0), \mathbf{0}_{1\times 6(k-s)}\right]'$. The following restrictions upon vectors of coefficients $\mathbf{a}_0$ and $\mathbf{a}_1$ can be derived from the Phillips curve above:

$$\begin{aligned}
\mathbf{a}_0 &= (1-\theta)(1-\beta\theta)\left[\boldsymbol{\nu}\mathbf{m}_1 - \left(\sum_{s=0}^{k-1}(1-\theta)^s\,\boldsymbol{\gamma}_a^{(s)\prime}\right)\right] \\
&+\beta\theta\boldsymbol{\nu}\mathbf{m}_2 + \beta\theta\left(\sum_{s=0}^{k-1}(1-\theta)^s\right)\mathbf{1}_1^T\mathbf{v}_1\mathbf{v}_0
\end{aligned} \tag{45}$$

$$\mathbf{a}_1 = (1-\theta)(1-\beta\theta)\left(\sum_{s=0}^{k-1}(1-\theta)^s\right)\mathbf{1}_2^T\mathbf{v}_1 + \beta\theta\left(\sum_{s=0}^{k-1}(1-\theta)^s\right)\mathbf{1}_1^T\mathbf{v}_1\mathbf{v}_1 \tag{46}$$

where I define:

$$\mathbf{m}_1 \equiv \begin{bmatrix} \left[\mathbf{1}_2^T \mathbf{v}_0 \mathbf{T}^{(1)}\right] \\ (1-\theta)\left[\mathbf{1}_2^T \mathbf{v}_0 \mathbf{T}^{(2)}\right] \\ (1-\theta)^2\left[\mathbf{1}_2^T \mathbf{v}_0 \mathbf{T}^{(3)}\right] \\ \vdots \\ (1-\theta)^{k-1}\left[\mathbf{1}_2^T \mathbf{v}_0 \mathbf{T}^{(k)}\right] \end{bmatrix}, \ \mathbf{m}_2 \equiv \begin{bmatrix} \left[\mathbf{1}_1^T \mathbf{v}_0 \mathbf{M}\mathbf{T}^{(1)}\right] \\ (1-\theta)\left[\mathbf{1}_1^T \mathbf{v}_0 \mathbf{M}\mathbf{T}^{(2)}\right] \\ (1-\theta)^2\left[\mathbf{1}_1^T \mathbf{v}_0 \mathbf{M}\mathbf{T}^{(3)}\right] \\ \vdots \\ (1-\theta)^{k-1}\left[\mathbf{1}_1^T \mathbf{v}_0 \mathbf{M}\mathbf{T}^{(k)}\right] \end{bmatrix},$$

$$\boldsymbol{\nu} \equiv \mathbf{1}_{1\times k}$$

Finally, the Taylor rule (14) implies that:

$$\begin{aligned} \mathbf{c}_0\varphi_{t|t}^{(0:k)} + \quad & \mathbf{c}_1\mathbf{s}_{t-1} = \rho_r\hat{R}_{t-1} + (1-\rho_r)\left[\phi_\pi\left(\mathbf{a}_0\varphi_{t|t}^{(0:k)} + \mathbf{a}_1\mathbf{s}_{t-1}\right) + (1-\phi_\pi)\,\mathbf{1}_3^T\varphi_{t|t}^{(0:k)}\right] \\ & + (1-\rho_r)\,\phi_y\left(\mathbf{b}_0\varphi_{t|t}^{(0:k)} + \mathbf{b}_1\mathbf{s}_{t-1} - \widehat{y}_{t-1}\right) + \mathbf{1}_4^T\varphi_{t|t}^{(0:k)} \end{aligned}$$

It is simple to see that the equation above translates into the following restrictions:

$$\mathbf{c}_0 = (1-\rho_r)\left[\phi_\pi\mathbf{a}_0 + (1-\phi_\pi)\,\mathbf{1}_3^T + \phi_y\mathbf{b}_0\right] + \mathbf{1}_4^T \tag{47}$$

$$\mathbf{c}_1 = \rho_r\mathbf{1}_3^T + (1-\rho_r)\left[\phi_\pi\mathbf{a}_1 + \phi_y\left(\mathbf{b}_1 - \mathbf{1}_2^T\right)\right] \tag{48}$$

Equations (43)-(48) are a system of non-linear equations in the coefficients $\mathbf{v}_0 \equiv \left[\mathbf{a}_0', \mathbf{b}_0', \mathbf{c}_0'\right]'$ and $\mathbf{v}_1 \equiv \left[\mathbf{a}_1', \mathbf{b}_1', \mathbf{c}_1'\right]'$. For any given set of parameter values and matrices $\mathbf{M}$ and $\mathbf{N}$ of coefficients for equation (15), the solution for this system of equations can be found by using one of the many non-linear equation solvers. This task never turn out to be computationally challenging. I find zeros of a system of non-linear equations through Newton's method, discussed in Judd (1998) (pp. 167-8).[23]

C Transition Equation of High–Order Expectations

In this section, we show how to derive the law of motion of the average higher-order expectations of the exogenous variables (i.e., z_t, $\eta_{a,t}$, $\hat{\pi}^*_{t|t}$, $\eta_{r,t}$, g_t, $\eta_{g,t}$) for given parameter values and vectors of coefficients $\mathbf{v}_0$ and $\mathbf{v}_1$.[24] A quick inspection of equation (15) reveals that such a law of motion is entirely determined by the matrices $\mathbf{M}$ and $\mathbf{N}$. Since the model is linear and all shocks are Gaussian, one can pin down these matrices (i.e., solve firms' signal-extraction problem) through the Kalman filter, which requires the specification of firms' state-space model. Firms' reduced-form state-space model (19)-(20) can be concisely cast as follows:

$$\mathbf{X}_t = \mathbf{W}\cdot\mathbf{X}_{t-1} + \mathbf{U}\cdot\boldsymbol{\varepsilon}_t \tag{49}$$

$$\mathbf{Z}_t = \mathbf{D}_1\mathbf{X}_t + \mathbf{D}_2\mathbf{X}_{t-1} + \mathbf{Q}e_{j,t} \tag{50}$$

[23] I find that non-linear solver fails to find a solution when the parameter ϕ_π is too small.

[24] Recall that the vectors $\mathbf{v}_0$ and $\mathbf{v}_1$ have been shown to depend on parameters and the matrices $\mathbf{M}$ and $\mathbf{N}$ (see Appendix B). In this Section, these vectors are treated as given.

where $\mathbf{W} \equiv \mathbf{A}^{-1}\mathbf{B}$ and $\mathbf{U} \equiv \mathbf{A}^{-1}\mathbf{C}$ and the other notation is defined in Section 3.8. Recall that we defined

$$\mathbf{A} \equiv \begin{bmatrix} \mathbb{I} & \mathbf{0} \\ -\mathbf{v}_0 & \mathbb{I} \end{bmatrix}, \ \mathbf{B} \equiv \begin{bmatrix} \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{v}_1 \end{bmatrix}, \ \ \mathbf{C} \equiv \begin{bmatrix} \mathbf{N} \\ \mathbf{0} \end{bmatrix}$$

where the matrices $\mathbf{v}_0$ and $\mathbf{v}_1$ have some known values. Note that since three signals in $\mathbf{Z}_t$ are endogenous, the state vector $\mathbf{X}_t$ must include the endogenous state variables $\mathbf{s}_t \equiv \left[\hat{\pi}_t, \hat{y}_t, \hat{R}_t\right]'$.

Solving firms' signal extraction problem requires applying the Kalman filter. Note that firms' observation equations (50) include lagged state variables. This slightly modified the usual Kalman formula for signal extraction. The correct formula has been provided by Nimark (2010). Here we propose a(n alternative) Bayesian derivation of this formula.

The initial period:

In period 0, we start with a prior distribution for the initial state $\mathbf{X}_0$. This prior is of the form $\mathbf{X}_{0|0} \backsim \mathcal{N}\left(\mathbf{X}_{0|0}, \mathbf{P}_{0|0}\right)$.

Forecasting:

At $(t-1)^+$, that is after observing Z_{t-1}, the belief about the state vector has the form $\mathbf{X}_t|\mathbf{Z}^{t-1} \backsim \mathcal{N}\left(\mathbf{X}_{t|t-1}, \mathbf{P}_{t|t-1}\right)$ where

$$\begin{aligned} \mathbf{X}_{t|t-1} &= \mathbf{W}\mathbf{X}_{t-1|t-1} \\ \mathbf{P}_{t|t-1} &= \mathbf{W}\mathbf{P}_{\mathbf{t-1|t-1}}\mathbf{W}' + \mathbf{U}\mathbf{U}' \end{aligned} \tag{51}$$

The density

$$\mathbf{Z}_t|\mathbf{Z}^{t-1} \backsim \mathcal{N}\left(\mathbf{Z}_{t|t-1}, \mathbf{F}_{t|t-1}\right)$$

where

$$\mathbf{Z}_{t|t-1} = \mathbf{D}_1\mathbf{X}_{t|t-1} + \mathbf{D}_2\mathbf{X}_{t-1|t-1}$$

and $\mathbf{F}_{t|t-1} = E\left[\mathbf{Z}_t\mathbf{Z}_t'|\mathbf{Z}^{t-1}\right]$ and hence:

$$\mathbf{F}_{t|t-1} = E\left[\left(\mathbf{D}_1\mathbf{X}_t + \mathbf{D}_2\mathbf{X}_{t-1} + \mathbf{Q}e_{j,t}\right)\left(\mathbf{D}_1\mathbf{X}_t + \mathbf{D}_2\mathbf{X}_{t-1} + \mathbf{Q}e_{j,t}\right)'|\mathbf{Z}^{t-1}\right]$$

and by working the product out, one obtains:

$$\begin{aligned} \mathbf{F}_{t|t-1} &= \mathbf{D}_1\mathbf{P}_{t|t-1}\mathbf{D}_1' + \mathbf{D}_2\mathbf{P}_{t-1|t-1}\mathbf{D}_2' + \mathbf{Q}\mathbf{Q}' + \\ &\quad + \mathbf{D}_1 E\left[\mathbf{X}_t\mathbf{X}_{t-1}'|\mathbf{Z}^{t-1}\right]\mathbf{D}_2' + \mathbf{D}_2 E\left[\mathbf{X}_{t-1}\mathbf{X}_t'|\mathbf{Z}^{t-1}\right]\mathbf{D}_1' \end{aligned} \tag{52}$$

Note that

$$\begin{aligned} E\left[\mathbf{X}_t\mathbf{X}_{t-1}'|\mathbf{Z}^{t-1}\right] &= E\left[\left(\mathbf{W}\cdot\mathbf{X}_{t-1} + \mathbf{U}\cdot\boldsymbol{\varepsilon}_t\right)\mathbf{X}_{t-1}'|\mathbf{Z}^{t-1}\right] \\ &= \mathbf{W}\mathbf{P}_{\mathbf{t-1t-1}} \end{aligned}$$

Combining this result with equation (52) yields:

$$\begin{aligned} \mathbf{F}_{t|t-1} &= \mathbf{D}_1\mathbf{P}_{t|t-1}\mathbf{D}_1' + \mathbf{D}_2\mathbf{P}_{t-1|t-1}\mathbf{D}_2' + \mathbf{Q}\mathbf{Q}' + \\ &\quad + \mathbf{D}_1\mathbf{W}\mathbf{P}_{\mathbf{t-1t-1}}\mathbf{D}_2' + \mathbf{D}_2\mathbf{P}_{t-1t-1}\mathbf{W}'\mathbf{D}_1' \end{aligned}$$

by substituting equation (51) into the equation above leads to

$$\begin{aligned} \mathbf{F}_{t|t-1} &= \mathbf{D}_1\mathbf{W}\mathbf{P}_{\mathbf{t-1|t-1}}\mathbf{W}'\mathbf{D}_1' + \mathbf{D}_2\mathbf{P}_{t-1|t-1}\mathbf{D}_2' + \mathbf{Q}\mathbf{Q}' + \\ &\quad + \mathbf{D}_1\mathbf{W}\mathbf{P}_{\mathbf{t-1t-1}}\mathbf{D}_2' + \mathbf{D}_2\mathbf{P}_{t-1t-1}\mathbf{W}'\mathbf{D}_1' + \mathbf{D}_1\mathbf{U}\mathbf{U}'\mathbf{D}_1' \end{aligned}$$

and finally to:

$$\mathbf{F}_{t|t-1} = (\mathbf{D}_1\mathbf{W} + \mathbf{D}_2)\,\mathbf{P}_{t-1|t-1}\,(\mathbf{D}_1\mathbf{W} + \mathbf{D}_2)' + \mathbf{Q}\mathbf{Q}' + \mathbf{D}_1\mathbf{U}\mathbf{U}'\mathbf{D}_1'$$

The joint distribution of $\mathbf{X}_t$ and $\mathbf{Z}_t$, thus, is

$$\begin{bmatrix} \mathbf{X}_t \\ \mathbf{Z}_t \end{bmatrix} |\mathbf{Z}^{t-1} \backsim \mathcal{N}\left(\begin{matrix} \mathbf{X}_{t|t-1} \\ \mathbf{Z}_{t|t-1} \end{matrix}, \begin{matrix} \mathbf{P}_{t|t-1} & \mathbf{P}_{t|t-1}\mathbf{D}_1' + \mathbf{W}\mathbf{P}_{t-1|t-1}\mathbf{D}_2' \\ \mathbf{D}_1\mathbf{P}_{t|t-1}' + \mathbf{D}_2\mathbf{P}_{t-1|t-1}\mathbf{W}' & \mathbf{F}_{t|t-1} \end{matrix} \right)$$

as

$$\begin{aligned} E\left[\mathbf{X}_t\mathbf{Z}_t'|\mathbf{Z}^{t-1}\right] &= \mathbf{P}_{t|t-1}\mathbf{D}_1' + E\left(\mathbf{X}_t\mathbf{X}_{t-1}'|\mathbf{Z}^{t-1}\right)\mathbf{D}_2' \\ &= \mathbf{P}_{t|t-1}\mathbf{D}_1' + E\left[\left(\mathbf{W}\cdot\mathbf{X}_{t-1} + \mathbf{U}\cdot\boldsymbol{\varepsilon}_t\right)\mathbf{X}_{t-1}'|\mathbf{Z}^{t-1}\right]\mathbf{D}_2' \\ &= \mathbf{P}_{t|t-1}\mathbf{D}_1' + \mathbf{W}\mathbf{P}_{t-1|t-1}\mathbf{D}_2' \end{aligned}$$

where the second equality follows from using equation (49). Hence, it follows that

$$\mathbf{X}_t|\left(\mathbf{Z}_t, \mathbf{Z}^{t-1}\right) = \mathbf{X}_t|\mathbf{Z}^t \backsim \mathcal{N}\left(\mathbf{X}_{t|t}, \mathbf{P}_{t|t}\right)$$

where[25]

$$\mathbf{X}_{t|t} = \mathbf{X}_{t|t-1} + \left[\mathbf{P}_{t|t-1}\mathbf{D}_1' + \mathbf{W}\mathbf{P}_{t-1|t-1}\mathbf{D}_2'\right]\mathbf{F}_{t|t-1}^{-1}\left[\mathbf{Z}_t - \mathbf{Z}_{t|t-1}\right] \tag{53}$$

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \left[\mathbf{P}_{t|t-1}\mathbf{D}_1' + \mathbf{W}\mathbf{P}_{t-1|t-1}\mathbf{D}_2'\right]\mathbf{F}_{t|t-1}^{-1}\left[\mathbf{D}_1\mathbf{P}_{t|t-1}' + \mathbf{D}_2\mathbf{P}_{t-1|t-1}\mathbf{W}'\right] \tag{54}$$

Therefore, combining equation (54) with equation (51) yields:

$$\mathbf{P}_{t+1|t} = \mathbf{W}\left[\mathbf{P}_{t|t-1} - \left(\mathbf{P}_{t|t-1}\mathbf{D}_1' + \mathbf{W}\mathbf{P}_{t-1|t-1}\mathbf{D}_2'\right)\mathbf{F}_{t|t-1}^{-1}\left(\mathbf{D}_1\mathbf{P}_{t|t-1}' + \mathbf{D}_2\mathbf{P}_{t-1|t-1}\mathbf{W}'\right)\right]\mathbf{W}' + \mathbf{U}\mathbf{U}' \tag{55}$$

Recall equation (50) and write the law of motion of the firm j's first-order beliefs about $\mathbf{X}_t$ as

$$\mathbf{X}_{t|t} = \mathbf{X}_{t|t-1} + \left[\mathbf{P}_{t|t-1}\mathbf{D}_1' + \mathbf{W}\mathbf{P}_{t-1|t-1}\mathbf{D}_2'\right]\mathbf{F}_{t|t-1}^{-1}\left[\mathbf{D}_1\mathbf{X}_t + \mathbf{D}_2\mathbf{X}_{t-1} + \mathbf{Q}e_{j,t} - \left(\mathbf{D}_1\mathbf{X}_{t|t-1} + \mathbf{D}_2\mathbf{X}_{t-1|t-1}\right)\right]$$

where we have combined equations (53) and (50). By recalling that $\mathbf{X}_{t|t-1}(j) = \mathbf{W}\mathbf{X}_{t-1|t-1}(j)$, we

[25] Here we use the following lemma to get the moments in the text. Let the random vector $(x', y')' \backsim \mathcal{N}(\mu, \Sigma)$ such that we denote

$$\mu = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix} \text{ and } \Sigma = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}$$

Then the pdf $x|y \backsim \mathcal{N}\left(\mu_{x|y}, \Sigma_{xx|y}\right)$ with

$$\begin{aligned} \mu_{x|y} &= \mu_x + \Sigma_{xy}\Sigma_{yy}^{-1}\left(y - \mu_y\right) \\ \Sigma_{xxy} &= \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\sigma_{yx} \end{aligned}$$

have:

$$\begin{aligned}
\mathbf{X}_{t|t} &= \mathbf{W}\mathbf{X}_{t-1|t-1} + \\
&+\underbrace{\left[\mathbf{P}_{t|t-1}\mathbf{D}_1' + \mathbf{W}\mathbf{P}_{t-1|t-1}\mathbf{D}_2'\right]\mathbf{F}_{t|t-1}^{-1}}_{\mathbf{K}}\left[\mathbf{D}_1\mathbf{X}_t + \mathbf{D}_2\mathbf{X}_{t-1} + \mathbf{Q}e_{j,t} - \left(\mathbf{D}_1\mathbf{W}\mathbf{X}_{t-1|t-1} + \mathbf{D}_2\mathbf{X}_{t-1|t-1}\right)\right]
\end{aligned}$$

By rearranging one obtains:

$$\mathbf{X}_{t|t} = \left(\mathbf{W} - \mathbf{K}\mathbf{D}_1\mathbf{W} - \mathbf{K}\mathbf{D}_2\right)\mathbf{X}_{t-1|t-1} + \mathbf{K}\left[\left(\mathbf{D}_1\mathbf{W} + \mathbf{D}_2\right)\cdot\mathbf{X}_{t-1} + \mathbf{D}_1\mathbf{U}\cdot\boldsymbol{\varepsilon}_t + \mathbf{Q}e_{j,t}\right]$$

The vector $\mathbf{X}_{t|t}$ contains firm j's first-order expectations about model's state variables. Integrating across firms yields the law of motion of the average expectation about $\mathbf{X}_t$:

$$\mathbf{X}_{t|t}^{(1)} = \left(\mathbf{W} - \mathbf{K}\mathbf{D}_1\mathbf{W} - \mathbf{K}\mathbf{D}_2\right)\mathbf{X}_{t-1|t-1}^{(1)} + \mathbf{K}\left[\left(\mathbf{D}_1\mathbf{W} + \mathbf{D}_2\right)\cdot\mathbf{X}_{t-1} + \mathbf{D}_1\mathbf{U}\cdot\boldsymbol{\varepsilon}_t\right]$$

Note that $\varphi_{t|t}^{(0:\infty)} = \left[\varphi_t, \varphi_{t|t}^{(1:\infty)}\right]'$ and that:

$$\varphi_t = \underbrace{\begin{bmatrix} \rho_z & 0 & 0 & 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & \rho_\pi & 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & 0 & 0 & \rho_g & 0 & \mathbf{0} \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} \end{bmatrix}}_{\mathbf{R}_1}\varphi_{t-1|t-1}^{(0:k)} + \underbrace{\begin{bmatrix} \sigma_z & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_a & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_\pi & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_r & 0 & 0 \\ 0 & 0 & 0 & 0 & \tilde{\sigma}_g & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_g \end{bmatrix}}_{\mathbf{R}_2}\cdot\boldsymbol{\varepsilon}_t$$

So by using the assumption of common knowledge in rationality, we can fully characterize the matrices $\mathbf{M}$ and $\mathbf{N}$:

$$\mathbf{M} = \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{6\times 6} & \mathbf{0}_{6\times 6k} \\ \mathbf{0}_{6k\times 6} & \left(\mathbf{W} - \mathbf{K}\mathbf{D}_1\mathbf{W} - \mathbf{K}\mathbf{D}_2\right)|_{(1:6k,1:6k)} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{K}\left(\mathbf{D}_1\mathbf{W} + \mathbf{D}_2\right)|_{(1:6k,1:6(k+1))} \end{bmatrix} \tag{56}$$

$$\mathbf{N} = \begin{bmatrix} \mathbf{R}_2 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{K}\mathbf{D}_1\mathbf{U}|_{(1:6k,1:6)} \end{bmatrix} \tag{57}$$

where $\cdot|_{(n_1:n_2,m_1:m_2)}$ denotes the submatrix obtained by taking the elements lying between the n_1-th row and the n_2-th row and between the m_1-th column and the m_2-th column. Note that $\mathbf{K}$ in the above equation denotes the steady-state Kalman gain matrix, which is obtained by iterating the equations (54)-(55) and the equation for the Kalman-gain matrix below:

$$\mathbf{K} = \left[\mathbf{P}_{t|t-1}\mathbf{D}_1' + \mathbf{W}\mathbf{P}_{t-1|t-1}\mathbf{D}_2'\right]\mathbf{F}_{t|t-1}^{-1}$$

until convergence. Equations (56)-(57) are equations (22)-(23) in the text.